

Herbert A. Simon and the concept of rationality: Boundaries and procedures

GUSTAVO BARROS*

This paper discusses Herbert A. Simon's conception of rationality in two of its principal general definitions: bounded rationality and procedural rationality. It argues that the latter is the one that better synthesizes the author's view about rational behavior and that the former fills mainly a critical function. They are complementarily used by Simon in this sense. In spite of that, it is argued that it is the low degree of specificity of the concept of bounded rationality one of the reasons for its relatively greater success.

Keywords: Herbert A. Simon; bounded rationality; procedural rationality.

JEL Classification: D01; B31; B52.

INTRODUCTION

Herbert A. Simon was the self-proclaimed, and proclaimed, "prophet of bounded rationality" (Simon, 1996, p. 250; and Sent, 1997, p. 323). In spite of the tone of such statement, it is beyond doubt that, in economics at least, the concept of bounded rationality is firmly associated with Simon's name, and conversely, and

* Gustavo Barros is Economics Ph.D. student at FEA-USP. E-mail: gb@usp.br. This paper is a sequel to the author's MA dissertation about Herbert A. Simon's economic work (Barros, 2004) done under supervision of Prof. Ana Maria Afonso Ferreira Bianchi and presented to FEA-USP (Faculdade de Economia, Administração e Ciências Contábeis da Universidade de São Paulo). That research counted on financial support of CAPES and FINE. This paper also benefited much from interaction with the research group Complex (Economia e Complexidade) at FEA, in particular the argument here proposed owes much to discussions with Prof. Jorge Eduardo de Castro Soromenho, and comments made by Prof. Gilberto Tadeu Lima on both the original work and a preliminary version of this paper allowed for important improvements. I would also like to thank the comments on the original work made by Edmundo Juarez Filho and comments by participants of the II Research Workshop on Institutions and Organizations (FEA-USP, 2007) on a preliminary version of the paper. Submitted: August 2007; Approved: March 2009.

that his authority is, time and again, employed to buttress the use of the concept that is being done, nowadays in a relatively more frequent fashion, by diverse strands in the field of economic science.

Although Simon was part of the chorus of critics to the concept of global rationality from the early beginnings of his career, a more precise conceptualization, and even the term *bounded rationality* date from the mid-fifties. According to Klaes and Sent (2005), in all likelihood, the term first appears in print in the 1957 book *Models of man, social and rational*. As far as I know of, this is indeed the precise date of “birth” of bounded rationality. According to Simon himself, the position was already clearly outlined in a paper not much antecedent, and which was included in the 1957 book (Simon, 1955; about this, see Silveira, 1994, p. 73). As a matter of fact, by this time, the concept had already the form and basic content it has up to today: the incapacity of exercise of global rationality makes the economic agents beings endowed with a bounded rationality. However, Simon, along his long lasting and prolific career, would advance much towards *specifying* his conception of rationality. One of the most important steps in this direction was the concept of *procedural rationality* (Simon, 1976b), proposed little before he was awarded the Nobel Prize, in 1978. To my judgement, procedural rationality has the ability to synthesize very adequately Simon’s view of rationality. Nevertheless, this second general concept of rationality has nowadays a much less marked presence in the economic science field and, as far as I know of, never had such a wide penetration as the concept of bounded rationality.¹

In this paper, I argue the following. First, that the concept of bounded rationality is characterized, above all, by its low degree of specificity. Second, that this characteristic can explain much of the (relative) current popularity of the concept. In the third place, I argue that Simon’s remaining main contributions to the debate on rationality and economic behavior, including the ones preceding 1976, can be grouped under the term “procedures” and, therefore, his behavioral theory is based on procedural rationality. In other words, it is the case of treating the concept of procedural rationality as the one that best expresses Simon’s view of rational behavior, to the detriment of bounded rationality, which mostly plays a critical role to mainstream economics. This paper is composed, besides this introduction, of four more sections. The second section discusses the concept of bounded rationality aiming at defining it and at pointing some of its important characteristics. The third section presents the concept of procedural rationality so that, in fourth section, we can discuss the relation between these two general concepts of rationality advanced by Simon. Special attention is dedicated to the historical chronology of the construction of these concepts by the author, bearing in mind that this chronological ordering helps in making clearer the logical relation between them. Finally, in the fifth section some final considerations are made.

¹ Klaes e Sent (2005, p. 42) point out that procedural rationality performed “a secondary role in his [Simon’s] corpus as well as in references to his work”.

BOUNDED RATIONALITY

Simon, throughout his work, consistently strived to build a theory of human behavior. This is true, in particular, with respect to his incursions in the economic field: what he produced was, above all, a theory of economic behavior. In other words, his focus was less “the economy” than “the economic agent”, though this agent does not necessarily equate to “individual”. Rationality is centrally placed in this behavioral theory: it is the main explaining element, although not the only one.

To Simon, the distance between rationality and behavior is bridged by the concept of “decision”.² A choice is a selection of one, among numerous possible behavior alternatives, to be carried out. Every behavior involves a selection of this kind, be it conscious or not. A decision is a process through which this selection is performed. Rationality is a criterion used in the decision that is theoretically grounded on the presupposition that the agents are *intendedly* rational. In other words, the agents *value* rationality as a criterion of choice and it is in this sense, and by this route, that rationality is taken as an explaining principle.

Rationality is defined by Simon as a relation of conformance (efficacy) between preestablished ends and the means to reach them. To him, the specification of these ends is a question of value and, hence, is beyond the scope of science. However, the relation between means and ends is a question of fact. The factual evaluation of this conformity involves, in theory, three “steps”: (i) the listing of all possible behavioral alternatives; (ii) the determination of all the consequences that will follow, in the future, to the adoption of each of these alternatives (in a determinist way or in the form of distributions of probabilities); (iii) the comparison of the alternatives, that should be evaluated by the sets of consequences following each one of them, according to the preestablished ends (utility, profit or any other specified pay-off function).

Up to this point, Simon does not distance from the canonic concepts of rationality, which appear under several names in his work: “global rationality”, “substantive rationality”, “the rationality of neoclassical theory”, “objective rationality”, “maximization”, “optimization”, “perfect rationality”, “strict rationality” and perhaps still others. In order to obtain some terminological homogeneity — even if at the risk of some imprecision — I’ll adopt the term “global rationality” to refer to all of them.

Simon, in *Administrative Behavior* (1947), though emphasizing the distinction between effective and theoretical behavior, assumes the model of global rationality. This point is worth stressing, for it is not usually appreciated in its proper dimen-

² Rigorously, “problem solving” should also be included here. However, in Simon’s definitions the frontier between “decision making” and “problem solving” is quite blurred: in general, the more important are search processes, the closer we will be of problem solving and the farther of decision making. I will treat both indistinctively by the term “decision”, for I understand this distinction, although important in other contexts, does not affect the main lines of argument here advanced.

sion, and leads to some very widespread misunderstandings concerning the history of the concept of bounded rationality. It is true that, already in 1947, the distance starts to appear in statements regarding the incapacity of the agents of carrying out, in practice, the “steps” listed above: (i) global rationality requires knowledge of all possible behavior alternatives, however just a few of these alternatives are considered; (ii) global rationality requires full knowledge and anticipation about all future consequences that will follow each alternative, however such knowledge is always very fragmentary; (iii) the valuation of the consequences has, too, to be “predicted”, and such prediction will depend, among other things, on imagination (1947, pp. 80-81). Such list separates the hypothesis of omniscience of the hypothesis of rationality: every item concerns the limits on the knowledge the agent effectively has, but that do not prevent the agent from acting rationally based on such knowledge. However, such a distance appears mainly as a means of establishing the limits of *the theory that is being used*. In other words, what Simon aims here is to establish the difference between *theoretical behavior* and *actual or practical behavior*.

A testimony of Antonio Maria da Silveira, who was Simon’s student at Carnegie Tech in the sixties and kept in touch with him afterwards, is particularly clarifying to this subject:

About Simon, the relevant to anticipate in this context is that he [...] started from the neoclassical illumination. Maximization was the theme of his classic book, *Administrative Behavior* (1947), in the same way as satisficing became the theme in another revolutionary classic, co-authored with March, *Organizations* (1958). [...] Simon verified in practice the direct inapplicability of neoclassical theory. It was too his commitment to the direct applicability of his theoretical work that drove him to the change, to the foundation of what is nowadays established as behavioral economics. I presented directly to Simon these first verifications in a talk in June 1991. He not only confirmed, but also indicated me the paper in which the satisficing concept already appeared well characterized, ‘A behavioral model of rational choice’ (1955) [...]. (Silveira, 1994, p. 73)³

Administrative Behavior’s own purpose and structure make the argument clearer. In the first place, there is a whole chapter aimed at analyzing the influence of the “criterion of efficiency” in decision — where such criterion is defined in the following way: “*The criterion of efficiency dictates that choice of alternatives which produces the largest result for the given application of resources*” (Simon, 1947, p. 179, emphasis in the original) —, that would latter completely disappear from his theory. In the second place, it is necessary to have in mind that Simon was writing

³ Translation from Portuguese by the author. Silveira continues and comments the Portuguese translation he proposed to the term satisficing: “a propósito, Simon gostou muito de minha tradução de *satisficing* para *satisfazimento*, palavra buscada no português arcaico, em vez de *satisfação*, como vem sendo divulgado no Brasil; ele também gastou tempo na busca da palavra inglesa apropriada” (1994, p. 73).

a thesis in political science, having as subject decision processes within administrative organizations. More specifically, he was advancing a *theory of administration*. And he was making use of economic theory for that, or else, he was *applying economic theory to administration* according to his specific perspective. It is true that such an application of the canonic economic theory to a different ambit demanded, to Simon, an analysis regarding its conditions of validity and it is also true that he did not dispense with other explaining factors, besides rationality, to administrative behavior. These concerns forced him to discuss the “area of rationality” and its “limits”. However, what is central to notice is that Simon’s intellectual effort was directed not towards revising economic theory, but towards applying it. And towards applying it to another field, stretching the theory’s scope, even when this move would exact some flexibility.

The balance Simon tries to reach in *Administrative Behavior* is to include economic man’s maximization as a value premise to a rational administrator, but *without rendering trivial the administrative activity*. That is the reason why it is necessary to flexibilize economic theory when it is applied to the administrative field: if this is not done, the administrative task becomes banal and a theory of administration becomes useless. Not to do it would be to “solve the problem” — the problem Simon himself posed — by declaring it does not exist (Cf. Simon, 1947, pp. 240-1). A conclusion we can take from all this is that it is an anachronism to attribute to *Administrative Behavior* the emergence of the concept of bounded rationality. The long introduction to the third edition, which was published in 1976, suffers from this anachronism, and is perhaps one of the factors responsible for the confusion concerning this matter. But that does not change the situation.

Nevertheless, some years later, Simon (1955, p. 241) would clearly introduce boundaries to rationality proper, and which would become decisive: besides limited access to the several kinds of information, limits on the computational capabilities step in. On synthesizing this point, the formulation is the following:

The alternative approach employed in these papers is based on what I shall call the *principle of bounded rationality*: *The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world — or even for a reasonable approximation to such objective rationality.* (Simon, 1957, p. 198, see also p. 202)

As far as is known, this quote constitutes the first appearance in print of the term “bounded rationality”. An important aspect of it, and that is worth stressing in it, is that the concept of bounded rationality is built as the negative of the concept of global rationality. The boundaries listed are some, but this is not the central point. The essential is that the concept of bounded rationality is intended to en-

compass the idea of the *practical impossibility of exercise of global rationality*.⁴ This carries two implications. The first is that Simon, now, is effectively directing his firing power against global rationality: he is questioning economic theory, and speaking to and public of economists.⁵ Moreover, the basis of the confrontation is precisely the lack of realism of the presuppositions sustaining global rationality, resulting in an impossibility of application *without any mediation* to practical situations. It is not incidental the fact that we can find in Simon's arguments, frequently in a central role, plenty of expressions such as: "in fact", "in practice", "in the real world", "in real life", "really", "realistically" etc. The second implication is that bounded rationality is not a logically autonomous concept. In its definition it is present, implicitly or explicitly, and it must be present, the concept of global rationality, or some of its variants. And it must be so precisely because of what the concept tries to capture: if it is defined as "impossibility of global rationality" or as the "negation of global rationality" then it is a *derivate concept*. In other words, from a logical standpoint, it is necessary do define global rationality *first*, and *then* proceed to discuss its inapplicability or to negate it. The boundedly rational agent is that one who is incapable of, in practice, exercising global rationality. It is this incapacity that justifies that the theory directs its attention to these boundaries and to the different ways through which the agents circumvent them. If this is the negation aimed by the concept in the 1950s, this seems to be consistent with the author's position in the end of the 1990s:

Global rationality, the rationality of neoclassical theory, assumes that the decision maker has a comprehensive, consistent utility function, knows all the alternatives that are available for choice, can compute the expected value of utility associated with each alternative, and chooses the alternative that maximizes expected utility. Bounded rationality, a rationality that is consistent with our knowledge of actual human choice behavior, assumes that the decision maker must search for alternatives, has egregiously incomplete and inaccurate knowledge about the consequences of actions, and chooses actions that are expected to be satisfactory (attain targets while satisfying constraints). (Simon, 1997, p. 17)

Early on, Simon grouped the bounds to rationality — very limited knowledge and cognitive limits — in what he called "psychological properties" of the agent, and stated the need for the empirical study of such properties as an appropriate foundation to a theory of rational behavior which purports to be predictive and

⁴ And not of its logical impossibility. Simon's arguments, regarding this matter, are distinct from the attacks on global rationality through logical computability considerations. He proposes the necessity of a practical computability, that is, one that is possible "in reasonable time" and not simply finite.

⁵ The 1955 paper was published in *The Quarterly Journal of Economics*.

descriptive, and even prescriptive or normative. The theory of global rationality, says he, operates based only on an “objective” description of the environment of decision — the “external constraints” —, the agent being fully depicted by his or her preferences. Simon claims for the explicit inclusion in the theory of other characteristics of the agent, such as the knowledge he or she effectively has and his or her cognitive capabilities — the “internal constraints”. In the 1950s, the attempts he made to deal with this theoretical impasse went in two main directions. In the first place, the lack of realism he pointed in the theory implied a need for the empirical study of how decision making is performed in practice. Field studies proper are not absent from Simon’s work, but they are certainly not very representative. The attempt to empirically study decision making process was done principally through the joint use of laboratory experiments, observing subjects in the process of decision making over relatively simple and standardized problem situations, and the computer simulation of models conceived based on such experiments. In the second place, Simon proposed a series of “simplifications”⁶ which would make the decision making process more tractable to the agent. No doubt, the most important of them is the *satisficing* hypothesis:

In these two essays [the papers of 1955 and 1956] the focus is upon ways of simplifying the choice problem to bring it within the power of human computation. [...] The key to the simplification of the choice process in both cases is the replacement of the goal of *maximizing* with the goal of *satisficing*, of finding a course of action that is ‘good enough’. I have tried, in these two essays, to show why this substitution is an essential step in the application of the principle of bounded rationality. (Simon, 1957, pp. 204-5)⁷

According to this hypothesis, decision makers, instead of trying to maximize values in a given choice, aim at *satisficing*: they search for alternatives that are good enough according to some pre-established criteria. The decision maker *optimizes* if he or she chooses an alternative that is *the best* one, as judged by a criterion that allows comparing all alternatives between themselves. The decision maker *satisfices* if he or she chooses an alternative that *attends or exceeds* a set of minimal acceptability criteria, if he or she chooses a satisfactory alternative, but one that is not necessarily the unique, nor the best. Optimization requires

⁶ For example, in Simon (1955) they appear under the name of “the essential simplifications”, in Simon (1956) as “simplification of the choice mechanisms”, in Simon (1957) as “simplification of the choice process” and in March and Simon (1958) as “simplified models”.

⁷ The concept of satisficing emerges clearly, in its content, in Simon (1955), the term appears shortly afterwards in Simon (1956, pp. 261, 270-1). Other statements of the definition can be found scattered throughout his work, in general without significant variations in its content, that is, the use he does of the concept is consistent through time. Some reference points are: Simon (1957, p. 205; 1976a, pp. xxix-xxx; 1987) and March and Simon (1958, pp. 140-141).

computation several orders of magnitude more complex than satisficing. In general, the satisficing hypothesis is accompanied by search processes, for alternatives as well as for new information (learning). Satisficing is also compatible with incomplete orderings of alternatives and with multiple criteria of choice. Other relevant simplifications advanced by Simon include: (i) the adoption of simplified models of reality; and (ii) the factoring of decisions in hierarchical chains of means and ends.

It is important to notice that in the 1955 paper the ideas of computational capacity and demand — the latter implicitly, under the “environment” which the agent faces — are clearly posed:

Broadly stated, the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms, including man, in the kinds of environments in which such organisms exist. (Simon, 1955, p. 241)

This quote is taken from the very paper in which the content of satisficing is first advanced, although the term only came about in Simon (1956). According to the author himself, this is also the paper economists most frequently chose for citation to refer to bounded rationality and satisficing (Simon, 1996, p. 165).⁸ From the perspective of the argument here proposed, it is worth emphasizing the following about this quote. The specter of global rationality is still markedly present in the formulation: the comparison between the two types of rationality is still structural to the theory. The general attitude of the paper denotes it too, for he first describes the model of global rationality and then, subsequently, proposes a set of “essential simplifications”.⁹ The idea of “simplification” presupposes something that becomes simpler, and the standard of comparison is precisely global rationality. Notwithstanding, these simplifications in the process of choice proposed by Simon indeed advance in the direction of specifying the concept of rationality used by the author and go beyond, on account of this, the strict argument of the bounds to rationality. I argue next that these specifications can be grouped under the concept of procedural rationality, advanced by Simon in 1976, being this, then, the most appropriate general concept to capture Simon’s *positive* definitions of rationality.

⁸ What is, actually, a very curious fact, considering none of the two *terms* — neither bounded rationality, nor satisficing — was present in the paper.

⁹ The paper that pairs this one (Simon, 1956) — together forming “the central core of the theory of choice” advanced in the book of 1957 — proceeds in very distinct manner, and treats rationality in a more positive fashion. However, it was published in the *Psychological Review* and it is, therefore, not surprising to find that it gives less weight to the concept of rationality cogent in the field of economic science.

PROCEDURAL RATIONALITY

These simplifications in the mechanisms of choice advanced by Simon — and that, as we have seen, appeared in the mid-fifties (Simon, 1955, 1956) — are elements that bore already more importance to the decision making *process* and that would, later, constitute the central components of the concept of procedural rationality. Initially, the question of computation appears somewhat muted under the idea of computational “capacity”, but it is present, as was pointed above. I say that the idea of “capacity” hinders the complete development of computational issues, because computation is something that has an important qualitative dimension, and fundamentally procedural: the process is embodied in the program, which describes *the way* computing is to be done. These computational concerns, therefore, appear in Simon’s work simultaneously to the concept of satisficing, initially labeled “satisfactory pay-offs”. Satisficing is essentially the hypothesis that allows, and practically induces, to the conception of diverse decision procedures. With it, the decision maker does not have to take into account all possible behavior alternatives and, in addition, does not need to worry about ascertaining whether the alternatives he or she is considering are, in fact, all the possible ones. Alternatives can be sequentially found out, by search processes, search being interrupted when a satisfactory alternative is found. Satisficing is, hence, the theoretical step that allows Simon to abandon the idea of rationality as a tautological reasoning over given premises, which permits rationality to operate in an open, not predetermined, space. On the other hand, satisficing forces him to inquire into the process by which such premises are built by the agent. The point I wish to emphasize is that, in the mid-fifties, although it is not yet the idea of procedure that organizes Simon’s efforts, the need to theorize about the decision procedure is already implied in his theoretical propositions. Moreover, and more importantly, starting from the critic of the boundaries to global rationality, every attempt at positive construction educates the procedural dimension of decision making. However, bounded rationality is always only the starting point and maintains its character of a construction in negative: “in conditions of bounded rationality” the agents resort to other expedients, different from those prescribed by global rationality, in order to exercise their intention of rationality. The specification of such expedients, of other types of rational behavior, is the reaction to a bounded rationality condition, but it is not bounded rationality itself.

The fact that these two elements — computation and satisficing — appear in Simon’s work in 1955 is not casual. In 1952, he became a consultant to RAND Corporation, initially involved in simulations of an air-defense early warning station, and then, from 1955 on, connected with the Computer Science Department. RAND was the paradigmatic military think tank in the post-Second World War period. It was also the world’s largest computational structure for scientific ends at the time. Simon’s entrance in RAND marks an intellectual inflection of his. Among the aspects of this change that interests us here is his distancing away from economics toward the areas of psychology and computer science, a move that

would only, and partly, be reverted in the 1970s — more specifically he placed himself in the nascent disciplines of cognitive psychology, cognitive science, artificial intelligence, operations research, and computer science, all of them tightly connected with the computer. His research program became essentially aimed at discovering *the symbolic processes that people use in thinking*, and was based on the exploration of an analogy between the computer and the human mind. The main method used was the combination of the tape-recording of the problem solving activity of subjects in the laboratory — producing “thinking-aloud protocols” — and of the simulation of computer programs that tried to emulate the activity registered in the laboratory. This meant that programs were taken to be theories: the program capable of simulating the human behavior recorded in the laboratory is, in itself, an explanation to that behavior. The attempt at programming (theorizing) the solution processes of relatively complex problems in computers with very limited memory and processing capacity led to the satisficing hypothesis, maximization would be impracticable without drastic simplification of the model. In other words, if, on the one hand, the mind-computer analogy suggests a very concrete image of what are the agents’ cognitive limits, on the other hand, programming always demands specification: what information the agent possesses, what criteria and procedures he or she uses to make decisions. Without such specifications, the programming cannot even begin.

It is based on his work at RAND and his contact with computers, then, that Simon starts to advance in a more positive manner other concepts of rationality, which diverged from global rationality. These would later (1976b) be grouped under the term “procedural rationality”, in an attempt at reinforcing the importance of the decision making process to the theory. Still concerning this matter, it is important to point that the very same basic theoretical elements that emerged in the 1950s as “simplifications” of the global rationality model form the core of the “procedures” in the 1970s, especially satisficing. Moreover, if the problems associated with computation were already in the fifties the main source of positive advances in the definition of rationality, they came to be central in the theory. More detailed comment upon these two issues is due.

Simon himself, by the late seventies, considered two concepts — already clearly present in his interventions in the economic science field in the fifties, and which we had the opportunity to discuss above — as the central elements to a more “positive” characterization of the mechanisms of decision: satisficing and search. In his own words:

In *Administrative Behavior*, bounded rationality is largely characterized as a residual category — rationality is bounded when it falls short of omniscience. And the failures of omniscience are largely failures of knowing all the alternatives, uncertainty about relevant exogenous events, and inability to calculate consequences. There was needed a more positive and formal characterization of the mechanisms of choice under

conditions of bounded rationality. [...] Two concepts are central to the characterization: *search* and *satisficing*. (Simon, 1979, p. 502)

The concepts of search and satisficing are intimately related. I have pointed above that it is the hypothesis of satisficing that allows for the relevance of search processes within decision making process. On the other hand, a mechanism of search — if it is not intended to be exhaustive, in which case it would be unnecessary to model it — needs a stop criterion, and Simon postulates satisficing for that, and points to empirical evidence sustaining that this is the criterion actually used by people in a wide range of situations, especially the more complex ones. Satisficing and search are, therefore, strongly complementary.

The second point in need of further comment is the one concerning computation. It has already been suggested that an important source of inspiration to the concept of satisficing, and to the use Simon does of search procedures in association with it, were his initial incursions in cognitive science, especially his attempts to program computers to imitate human decision making procedures and problem solving activity. The analogy between the human mind and the computer, in general, is taken in quite a literal fashion. One aspect of the theory that underlines this clearly is the practical identification between “computation” and “rational procedures” — people compute, computers think etc.

Economics, says Simon, “has to be concerned with computation”, with “the processes people actually use to make decisions”. However, these processes are precisely the object of another discipline: cognitive science. Economics is therefore, in this sense, tributary to cognitive science. It seems to me clear enough that the origin of Simon’s formulations about rationality is, from the mid-fifties on, cognitive science. His intervention in economics is fully coherent with his work in that area.

In defining procedural rationality, Simon (1976b) defines also another concept as counterpoint, substantive rationality. Behavior is substantively rational when it is adequate to the realization of given ends, subject to given conditions and constraints. Behavior is procedurally rational when it is the outcome of appropriate deliberation. Global rationality is understood as substantive in the sense that it is only concerned with *what* is the choice done, with its *result*. The concept of procedural rationality focuses on *how* the choice is done. The crucial issue in the distinction between substantive and procedural rationality lies in the proposition that the decision making process, and therefore, also the agent that carries out this process, influences crucially the decision result. Simon’s research in the area of cognitive science, demonstrated that, in complex situations, the choice taken, its result, strongly depended on the particular process that generated it, and not only on the objectives that oriented it. Hence, it becomes indispensable to know the *process* by which the choice is taken. As we have seen, this is precisely what Simon had been doing — more or less explicitly and consciously — since the 1950s.

We have also already noticed the close relation that exists between “decision

procedures” and “computation”. What I expect to be clear at this point is that the central question regarding procedural rationality is computational: procedures are *algorithms*. Simon conceived satisficing and search processes as algorithms, since they were forms of practical implementation (programming) of decision procedures in the computer. Moreover, it is worth emphasizing that these concepts, at least in their publication, historically *preceded* the term “bounded rationality”. The idea of choice of satisfactory alternatives emerges in Simon (1955), the term “satisficing” associated with search for alternatives appears fully developed in Simon (1956), the term “bounded rationality” is the first general concept that tries to encompass those simplifications/procedures and appears in Simon (1957). The second general concept, which came later, and that attempts to embrace *the very same mechanisms* is “procedural rationality”, which appears in Simon (1976b).¹⁰ In his autobiography, Simon (1996) comments the 1955 paper — “mostly written in 1952 during my first RAND summer” — in the following way:

What made the paper distinct from most contemporary economic writing was its explicit concern for the process of making decisions, for procedural and not just substantive rationality. Because of this concern with process, the paper also represents a first step toward computer simulation of human behavior. (Simon, 1996, pp. 165-6)¹¹

Summing up, the way by which Simon models rational behavior is, since very early, founded on procedures, the basis of which is composed by satisficing and by search processes. In this sense, and although it is an *a posteriori* imputation, the concept of procedural rationality is the one that best captures Simon’s view about rationality, as positively defined. The concept of bounded rationality, in its turn, tends always to operate by negation: the negation of global rationality. This argument could be questioned by saying that the problem is, at the bottom, just terminological, and that the concepts of bounded rationality and procedural rationality are really no more than two ways to look at the same thing, two points of view about the same set of theoretical principles. I would not oppose to it as a first approximation. However, to stop there implies, in my opinion, to lose something of what Simon has to tell us about rationality, and also to attribute to him more than what he has really done. A clear expression of the distinction I am delineating appears in the differences in reception of Simon’s rationality concepts: the repercussion of bounded rationality in economic science is much superior to the one of its younger and hard working sister.

¹⁰ This is, as far as I know of, the first appearance of the term *in Simon’s work*.

¹¹ The fact that this comment is done, evidently, in hindsight, does not affect the argument here proposed, once it is also, and explicitly, an imputation.

RATIONALITIES, BOUNDARIES AND PROCEDURES

Even after having proposed the concept of procedural rationality, in 1976, Simon continued to privilege bounded rationality as the main piece in his arguments. An example, among many possible ones, can be found in his Nobel Lecture, where he affirms that the results of his research in cognitive psychology supplied “rather conclusive empirical evidence that the decision-making process in problem situations conforms closely to the models of *bounded rationality*” (Simon, 1979, p. 507, emphasis added). Moreover, Simon in many instances practically equates “models of bounded rationality” with models that assume satisficing instead of maximization (for example, the quote above, of 1997, p. 17). One way to conduct this issue is to assume that the relation between the concepts of bounded rationality and procedural rationality is always one of *compatibility*, but not one of identity. I don’t believe that Simon himself would be comfortable with this proposition, however, not to recognize this use that he makes of the concepts implies a problem: if procedural rationality is to be considered an “advance” over bounded rationality, why then was not bounded rationality abandoned by him in favor of procedural rationality? No doubt, he continues to use them both parallelly, and, in general, bounded rationality constitutes the public and most visible face of Simon’s conception of rationality. We could say, alternatively, that procedural rationality was a frustrated attempt, from the point of view of its repercussion. Notwithstanding, to recognize the *complementarity* of the concepts seems to be the most appropriate solution to the question: bounded rationality does the critical part of the work while procedural rationality does the assertive one. An alternative formulation to this complementarity is to say that “under *conditions* of bounded rationality” a “more positive and formal characterization of the *mechanisms* of choice” is needed (Simon, 1979, p. 502, emphasis added), or else, a specification of the decision *procedures*. One quite rare instance of recognition of the difference, in the sense I am emphasizing, can be found in the following quote:

That case [the case of bounded rationality], at least as presented in the economics literature, had been a largely negative one, an attack on the veridicality of neoclassical theory without much more than hints about how to replace it. This distinction between procedural and substantive rationality, which I then began to develop, provided an opportunity to sketch out positively the (psychological) theory of procedural rationality. (Simon, 1996, p. 324)

However, a certain ambiguity results from this treatment dispensed by Simon to the concepts. At times bounded rationality is, or should be, understood as a negation of global rationality, and no more than that. At other times, it should be understood as a positive construction, which includes satisficing and search processes, a content which, as I argue here, would be better expressed by the term

“procedural rationality”. Simon himself does not usually put much effort into marking the distinction.

The result of this situation is that the concept, once it gained course in economic science, serves as a convenient shortcut to any models that refuse global rationality, and not necessarily those that Simon had in mind. Of course that this, in itself, does not constitute a problem neither to him nor to those who use the concept. What is interesting to point is that, if bounded rationality is indeed a frontal attack to global rationality theories, it stands out for its *lack of specificity*. This is true in Simon himself, but becomes especially evident when others adopt bounded rationality with positive rationality concepts distinct from Simon’s. What I suggest is that it is perhaps precisely this characteristic that enables bounded rationality to embrace much of the diversity of the so-called “heterodoxy” — and even something of the “orthodoxy”. When Simon compiled his economic papers, in the early 1980s, he entitled the two resulting volumes *Models of bounded rationality*: they are therefore “models” of bounded rationality, *some* models, and not “the models” and even less “the model”. Plurality is implicit in the concept. To bear this in mind makes easier to understand the use of the concept of bounded rationality by a Thomas Sargent, and the differences in the interpretations of this concept between Simon and Sargent (Sent, 1997; see also Sent, 2005). Klaes and Sent, studying that which they defined as the “bounded rationality’s semantic field”, follow historically the diverse expressions that denote the boundaries or limits to rationality, and also the different uses of some of the most important of these expressions. Based on this study, they formulate precisely the point in question.

It is thus an important aspect of the more recent use of ‘bounded rationality’ subsequent to its institutionalization as the core of the BR field that an increasing number of literatures began to use it in ways not only incongruent with the initial motivation of Simon when he crafted it, but also exhibiting significant cross-sectional divergence in interpretation. As we write, ‘bounded rationality’ is being employed with numerous different shades of meaning, and there is little indication of any convergence toward a dominant interpretation. All this has done little harm to the use of the expression as the main currency for conceptualizing limitations to the decision-making capabilities of human actors. (Klaes and Sent, 2005, p. 49)

This sets the stage for us to deal with another problem. Simon was not the first, and neither the only, to question the economic theory based on global rationality for its lack of realism. It is not difficult to suppose that this critic is as old as the theory. However, according to Klaes and Sent (2005, p. 45), by the late 1980s, bounded rationality “was firmly entrenched as one of the core concepts of economics, documented by its appearance in the main professional dictionary of the discipline of economics”. In this case, why then was Simon better succeeded than the

others were? (If he was) Why did he become one of the main spokespersons of this critic? (That, no doubt, he was.) Why did he become the “prophet” of bounded rationality?

Some (non excluding) hypothesis can be raised on this respect. In the first place, Simon confronts the theories of global rationality, it is true, but in their own field. There is common ground between his theoretical propositions and the more orthodox streams of economics: we should not disregard that rationality is the basic explanative element, that the economic agent is the *locus* of this rationality, and that economic modeling should take, preferably, formal mathematical shape. In the second place, he had far from negligible social and political insertion in the economic science field. Simon himself explains the Nobel he received this way: “if I was an outsider to the economics profession as a whole, I was an insider to its elite. Without that accreditation, I suspect I would not have won the prize.” (Simon, 1996, p. 326). By economics profession elite he meant Cowles Commission and, especially, the Econometric Society. In the third place, the Nobel Prize itself, received by him in 1978, doubtless weights in the legitimacy attributed to his work. This hypothesis gains some strength when we look the graphic elaborated by Klaes and Sent (2005, p. 39) registering the number of occurrences of the diverse expressions that compose the “semantic field of bounded rationality”. In it, we notice certain equilibrium between the different expressions up to the year 1975 and a clear “take off” of “bounded rationality” between 1975 and 1980. In other words, the Swedish academy’s influence on the success of the expression “bounded rationality” is, no doubt, significant. This is an interesting fact, concerning our general argument, for still another reason: the “take off” occurs when procedural rationality had already been brought up.¹² And last, but no less important, he himself offered, throughout his career, a series of specifications of the concept of rationality, actually at least since bounded rationality was proposed, through models defined in more positive fashion — although without pretension that theses propositions corresponded to the totality of the concept. This certainly helped to turn it more operational.

FINAL CONSIDERATIONS

We have discussed in this paper Herbert A. Simon’s conception of rationality. The purpose was to make explicit the relation between the two general rationality concepts of the author, and their respective contents. Bounded rationality and procedural rationality are used by him as complementaries. Bounded rationality is essentially a construction in negative: it is the negation of global rationality. It is

¹² What makes more difficult to simply sustain that “procedural rationality” was to perform a secondary role in his [Simon’s] corpus as well as in references to his work” because, by 1978, “Simon had embraced and become known for his insights concerning the concept of bounded rationality” as do Klaes and Sent (2005, p. 42). Although this fact certainly plays a role here.

marked, above all, by its low degree of specificity. The second one embodies, from a theoretical point of view, Simon's positive contributions to the concept of rationality. We can say that procedural rationality is the set of specifications — notably satisficing, but always in the form of procedures — proposed by Simon regarding what is rationality. Procedural rationality is the concept which best synthesizes Simon's view about rational behavior. The two concepts are complementary, then, in the following sense: bounded rationality does the critical part, and procedural rationality, the assertive one.

I have also proposed the hypothesis according to which it is the lack of specificity of bounded rationality one of the reasons why it finds greater resonance in the economic science field than procedural rationality. The differentiated reception of the two concepts, and the manner in which they are used, by Simon himself and by others, points in this direction. In short, I expect to have demonstrated that there is more in Herbert Simon than bounded rationality. However, I also suggest that there is less in bounded rationality than is customary to admit, and that such concept approaches very closely the critic, already quite ancient and diffused, of the lack of realism of the conventional theory's presuppositions, with a particular garment. Moreover, the concept of bounded rationality, in Simon's acception or in other's, represents an ambiguous move concerning the value attributed to the hypothesis of rationality in economic theory. On the one hand, and this Simon himself makes explicit, bounded rationality broadens the scope of the concept, in the sense that a greater set of economic situations can be treated as rational, presumably more realistically too. On the other hand, and this he does not mention, bounded rationality implies, in practice, a loss of specificity of the concept of rationality. As we have seen, when working with it, we are continuously entreated to specify, as Simon was. And the idea of rationality starts to appear in the plural — procedures, rationalities — where each of these "rationalities" is necessarily circumstantiated. Under these conditions, rationality sees potentially threatened its position of economic theory's explaining factor *par excellence*. Moreover, the systematic polysemy of the concept leaves room to several interpretations, not necessarily compatible among themselves: the bounded rationality cocoon can conceivably nest the most diverse metamorphosis of economic theory. This, as should be evident, does not rule out the pertinence of the critic that such concept operates and gives voice to. In addition, even this lack of specificity of the concept has non-negligible positive implications, in the sense that it is potentially creative, and provides space for theoretical innovation. I am inclined, as I argued above, to take such multiplicity of meanings to be found in the use of bounded rationality as a result of its unspecific character. We have here a case, then, of potentially creative destruction. It remains to be seen what precisely is being created.

Another important question regarding these characteristics — lack of specificity and use in multiple meanings — of bounded rationality is that they are predictably not long lasting, and they shall be less perennial the greater the concept's success and penetration. If today these characteristics of the concept permit, in the politics of

dispute among theories, bounded rationality to serve as a single and relatively compact banner to a relatively heterogeneous group, once the concept enlarges is course, and the value of being associated to it grows, internal disputes will eventually prevail and, in this case, the center of the discussion about bounded rationality will shift to the dispute around what are its “fundamental”, “original”, “canonic”, or “true” characteristics. In this regard, Simon will certainly have a privileged position, but such dispute, if it comes to happen, will most likely include the participation of the presently hegemonic stream, which is already being capable of absorbing bounded rationality and of finding even some functionality in it, as is well demonstrated in Sent’s works about the uses of bounded rationality by such figures as Thomas J. Sargent, Robert Aumann, and Kenneth Arrow (Sent, 2005). Anyhow, Simon himself, besides “prophesying”, proposed his version: procedural rationality, especially in the form of satisficing and search processes. According to him, “a theory of bounded rationality is necessarily a theory of procedural rationality” (Simon, 1997, p. 19). However, though he has taken the lead, others have proposed and continue to propose their own versions, and such consensus as there appears to be around bounded rationality is, as we have seen, only very superficial.

REFERENCES

- BARROS, Gustavo (2004) *Racionalidade e organizações: Um estudo sobre comportamento econômico na obra de Herbert A. Simon*, MA dissertation, São Paulo: FEA-USP.
- KLAES, Matthias; Sent, Esther-Mirjam (2005) “A conceptual history of the emergence of bounded rationality”, *History of Political Economy*, 37: 1: 27-59.
- MARCH, James G. & Simon, Herbert A. (1958) *Organizations*, New York: John Wiley and Sons.
- SENT, Esther-Mirjam (1997) “Sargent versus Simon: bounded rationality unbound”, *Cambridge Journal of Economics*, 21: 323-338.
- SENT, Esther-Mirjam (2005) “Simplifying Herbert Simon”, *History of Political Economy*, 37(2): 227-232.
- SILVEIRA, Antonio Maria da (1994) “Aplicabilidade de teorias: microneoclássica e estratégia empresarial”, *Revista de Economia Política*, vol. 14, n. 2 (54), abril-junho: 53-76.
- SIMON, Herbert A. (1947) *Administrative behavior: A Study of Decision-making Processes in Administrative Organization*, 1st ed., New York: The Macmillan Company.
- SIMON, Herbert A. (1955) “A behavioral model of rational choice”, *The Quarterly Journal of Economics*, vol. 69, n. 1, February: 99-118, compiled in, and quoted from, Simon (1957: 241-260).
- SIMON, Herbert A. (1956) “Rational choice and the structure of the environment”, *Psychological Review*, vol. 63, March, compiled in, and quoted from, de, Simon (1957: 261-273).
- SIMON, Herbert A. (1957) *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*, New York: John Wiley and Sons.
- SIMON, Herbert A. (1976a) *Administrative Behavior: A study of Decision-making Processes in Administrative Organization*, 3rd ed., with new introduction, New York: The Free Press.
- SIMON, Herbert A. (1976b) “From substantive to procedural rationality”, in Spiro J. Latsis, *Method and Appraisal in Economics*, Cambridge: Cambridge University Press: 129-148.
- SIMON, Herbert A. (1979) “Rational decision making in business organizations” [Nobel Memorial Lecture], *American Economic Review*, 69(4), September: 493-513.

- SIMON, Herbert A. (1987) "Satisficing", in John Eatwell; Murray Milgate; e Peter Newman (eds.), *The new Palgrave: A dictionary of economics*, vol. 4, New York: Palgrave: 243-245.
- SIMON, Herbert A. (1996) *Models of My Life*, Cambridge, MA: The MIT Press, 1996, originally published: New York: Basic Books, 1991.
- SIMON, Herbert A. (1997) *An Empirically Based Microeconomics*, Cambridge, UK: Cambridge University Press.