

**BOUNDED AND ADAPTIVE RATIONALITY, DECISION PROCESSES AND
PROBLEM SOLVING IN H. A. SIMON**

***RACIONALIDADE LIMITADA E ADAPTATIVA, PROCESSOS DE DECISÃO E
SOLUÇÃO DE PROBLEMAS EM H. A. SIMON***

Carolina Leister

carolina.Leister@fgv.br

EESP-FGV

J.R.N. Chiappin

chiappin@usp.br

FEA-USP

Abstract: *The purpose of this paper, which belongs to a series of papers we developed relative to Simon's theories of rationality, decision, company and organization, is to rationally reconstruct his seminal work on adaptive rationality and decision. The field of study of rationality and decision theories can be divided into two research programs: (i) the normative program, which interprets rationality as a set of axioms that supplies a logical structure for preferences and an optimization principle for the selection of alternatives; (ii) the descriptive program, according to which rationality is empirically reconstructed by means of the assumption that the decision-making individual incorporates a portfolio of adaptive heuristic strategies which operate satisfactorily, but not optimally, in the delineation of the opportunity set and in the selection of alternatives. In this paper we defend the thesis that Simon aims to confer greater empirical content to rationality, reinterpreting it as bounded and adaptive without, however, giving up rationality as a regulating idea.*

Keywords: *Bounded rationality. Adaptive rationality. Decision processes. Problem-solving approach. Heuristic strategies.*

Resumo: O objetivo deste artigo, que pertence a uma sequência de artigos desenvolvidos por nós e relativos às teorias da racionalidade, decisão, firma e organização de Simon, é reconstruir racionalmente seu trabalho seminal relativo à racionalidade e decisão adaptativas. O âmbito de estudo das teorias da racionalidade e decisão pode ser decomposto em dois programas de pesquisa: (i) o programa normativo, que interpreta racionalidade como um conjunto de axiomas que fornece uma estrutura lógica sobre as preferências e um princípio otimizador na seleção da alternativa; (ii) o programa descritivo, segundo o qual a racionalidade é reconstruída empiricamente por meio da suposição de que o indivíduo da decisão incorpora um portfólio de estratégias heurísticas adaptativas que operam satisfatória, mas não otimamente, no delineamento do conjunto de oportunidade e na seleção da alternativa. Neste artigo defendemos a tese de que Simon visa auferir maior conteúdo empírico à racionalidade, reinterpretando-a como limitada e adaptativa, sem, contudo, abrir mão da racionalidade como ideia reguladora em sua abordagem.

Palavras-chave: Racionalidade limitada. Racionalidade adaptativa. Processos decisórios. Abordagem de solução de problemas. Estratégias heurísticas.

Introduction

Rationality and decision theories are present in the theoretical core of microeconomics as a whole, and particularly in consumer theory. In the classical conception of rationality and decision, adopted by the economic mainstream, decision is considered to be the result of a rational procedure whose conditions are: (i) stable preferences about the alternatives of the opportunity set; (ii) complete information about the opportunity set; (iii) a consistent arrangement of the alternatives according to the decision-maker's preferences; (iv) a consistent attribution of the probabilities to states of the world in the case of decisions involving risk or uncertainty; (v) a maximization principle. Most commonly, however, these conditions are scarcely met. Decision psychologists, concerned with the construction of an empirically supported decision theory, turn their efforts to the empirical test of these assertions and observe that: (i) consumer preferences are not stable and vary according to the conditions of the problem, for instance, structuring their preferences differently in loss or gain contexts, as Tversky and Kahneman do in the prospects theory; (ii) consumers do not carry out an exhaustive research to define the opportunity set due to the costs incurred in this research process; (iii) the arrangement of the alternatives according to consumer preferences goes against the axioms that guarantee the consistence of the arrangement; (iv) the probabilities attributed by consumers/investors to possible states of the world do not satisfy the properties imputed to probability distributions; (v) consumers/investors do not always behave as maximizers.

If the conditions imputed to decision by the classical conception are maintained, the former can only be categorized as a normative decision-making proposition. What we mean is that it merely provides the conditions that *should* be met in order to obtain an optimal result, and not a conception that examines how consumers/investors really make their decisions. Even though this perspective of rationality is classified today as normative, in its original proposition it was admitted as an empirical feature of the decision-maker, and therefore as a property imputed to the model of the rational individual, and not as a normative and regulating idea. (In other words, it was supposed that the conditions enumerated above were fully met by the model of the rational individual.)

Seeking to elaborate an empirically feasible approach to rationality and decision, or descriptive as to the manner in which consumers/investors make decisions, a group of thinkers and experimentators gathered to construct the so-called descriptive conception of decision (Bell, Raiffa and Tversky, 1988; Leister, 2001). Even if many of these thinkers may be said to head this project, Simon is one of the first to propose and integrate a well-structured framework of decision and rationality, expressing his dissatisfaction with the program of rationality and normative decision. Our main thesis in this paper is that the Simonian proposition seeks to empirically reconstruct the rationality imputed to the decision-maker – decision being one of the categories of problems faced by his general problem solver. For Simon, rationality is as much a property of the individual as a regulating idea in the construction of his approach to rationality and decision. What we mean is that it is as much a property incorporated into his model of the individual as a normative restriction adopted in the construction of his approach.

It is because he reinterprets, on one hand, the conception of rationality as bounded and adaptive rationality and, on the other hand, does not give up rationality as a regulating idea

that we indicated Simon as the head of the descriptive research program of rationality and decision, instead of, say, Tversky and Kahneman. The reason for that is that these two seem more concerned with the construction of an empirical theory of decision, but are not necessarily committed to the preservation of rationality as a regulating idea in the construction of that theory. It is also for this reason that Simon's approach can be classified as a middle ground between, on one hand, the classical or normative conception of rationality and decision and, on the other hand, the descriptive conception, also known as prospects theory, by Tversky and Kahneman. Another of our theses concerning Simon's descriptive program is that he does not only elaborate a well-articulated theoretical framework but also, and foremost, operationalizes this framework in terms of a methodology for the construction and test of empirical theories, recovering the method of discovery as one of the central concerns of his metamethodological conception¹. However, this thesis will only be developed in a next paper of ours.

In the rational reconstruction that we shall carry out here we will make use of two methodological instruments of theoretical analysis, namely: (i) the research program apparatus (Lakatos, 1979); (ii) a version of the methodology of the theory of science (Chiappin, 1996). According to Lakatos, the research program is taken as the basic unit of the organization of knowledge and is characterized by a sequence of theories containing a theoretical core, with presuppositions that are common to its theories, and a heuristic, with the main instruments constructed by the different theories and made available for the solution of theoretical and empirical problems that appear inside the program. In this sense, we made reference to two programs of rationality and decision, the normative program and the descriptive program, each one disposing of a theoretical core and of its own positive heuristicⁱⁱ. The version of the methodology of the theory of science adopted in this paper identifies various theses within the core, among which are those that will be examined here: ontological theses, *i.e.*, theses concerning the basic agents of the theory. In Simon's case, they are the theses incorporated into his model of the rational individual which identify his or her main properties, and the theses involving a system of values and ends which operate as restrictions in the construction of the rationality and decision theories of the descriptive program, *i.e.*, restrictions that must be observed by theory. The main ones are the thesis that asserts Simon's commitment not to give up rationality as a regulating idea in his theory and the thesis which assumes that the empirical adequation of his model of the rational individual is a fundamental concern.

1. The Descriptive Research Program of Rationality and Decision and the Criticism of the Normative Program

In Simon's criticism of the normative program of rationality and decision he chooses as his main target the assumption of the model of the rational individual as an artificial construct, which does not permit an adequate portrayal of the empirical behavior of consumers and investors in their decisions. The requirement of a model of a decision-maker with perfect capacities derives from the rationality imputed to him or her by the classical conception, a perfect rationality in the sense that it should maximize the result of the decision and uphold the consistence and structure of their preferences about the opportunity set. Thus interpreted, the classical conception can only appear as a normative approach to rationality and decision and cannot describe how consumers/investors really decide. What happens, however, is that for many of the classical thinkers this rationality is supposed to be an empirical property which could be identified in empirical individuals, and not a regulating idea. So, it is because of the definition of rationality as a perfect rationality that the model of

the decision-maker is supposed to be a frail foundation on which the normative program of rationality and decision rests. To quote Simon:

If we closely examine the “classical” concepts of rationality..., we see immediately what severe demands they make upon the choosing organism (...). My first empirical proposition is that there is a complete lack of evidence that, in actual human choice situations of any complexity, these computations can be, or are in fact, performed (1979, p. 10).

Simon assumes, therefore, that the propositions which characterize the model of the rational individual in the normative conception are in need of an urgent revision. This is the point from which he erects his proposition and, in the course of this endeavor, builds the bedrock upon which a new research program of rationality and decision will be organized, called the descriptive program of rationality and decision. This conception, which intends to feature a strong empirical basis, leads to divergences with the normative perspective and generates a revision of the model of the rational individual in the normative program, and hence a redefinition of the notion of rationality. Thus, due to the emphasis put on the model of the rational individual and to the concern about its empirical adequation, Simon’s approach diverges from the normative program.

Given its estrangement in relation to the normative proposition and the purpose of elaborating an empirically based approach to rationality and decision, Simon’s program introduces in its model of the rational individual a psychology of decision. However, this psychology is not restricted to decision situations, but considers a series of problem-solving activities where decision appears only as one of the categories of those activities. The studies related to the subject of problem solving are done within cognitive sciences. Their object is the investigation of the human capacity for knowledge, and therefore a true gnoseology that involves a context of multidisciplinary researches, among which we may quote the philosophy of mind, neuroscience, cognitive and evolutionary psychology, artificial intelligence, linguistics and anthropology. In this sense, understanding Simon’s approach to decision entails the consideration of his psychological or, more specifically, cognitive theory and of his model of the rational individual.

His psychological proposition also contrasts with the behaviorist model idealized by Watson and Skinner, whose maximum requirement was the attempt to avoid at all cost the internalization of the behavioral determinants inside the mind, which would only produce animistic explanations of behavior (Pozo: 1998; Sternberg: 2000). To quote Skinner:

An even more common habit is to explain behavior in terms of an interior agent without physical dimensions, called “mental” or “psychic”. (...) In all those examples it is obvious that “mind” and “idea”, with their special characteristics, were invented ad hoc to supply spurious explanations. A science of behavior cannot expect much from these procedures. Given that mental or psychic events, it is said, do not have the characteristic dimensions of physical sciences, there is an additional reason to reject them (1989: p. 39-41).

In the behaviorist project behavior is treated exclusively as an answer to environmental determinants, the stimuli or conditioning factors of behavior. Also because of the behaviorist concern of producing empirically based theories, its approach ends up adopting an ultra-empiricist posture. Simon, on the contrary, adopts a less radical empiricist attitude, positing psychic determinants of behavior but making them feasible through an analogy between the cognitive processes of the human mind and computer processing. The

latter is then adopted as a tool to simulate and validate the assertions related to the processes of the former.

Hence, due to the positions defended within the Simonian descriptive research program of rationality and decision, we can say that, because of its empirical aspirations, it diverges, on one hand, from the normative perspective because of the extreme artificiality and apriorism of the latter in the construction of its model of the rational individual and, on the other hand, avoids slipping into behaviorism because of the latter's pure empiricism. In other words, the descriptive program of rationality and decision adopts an intermediate position between the classical or normative conception of rationality and decision and radical behaviorism, which is an ultra-empiricist conception of human behavior. Its main divergence is the assumption of a psychic and empirical model of the rational individual, absent from both approaches: from the first one because its model of the rational individual is not empirical but artificially constructed, from the second one because it takes away from it all its psychic content. Simon's attitude justifies, therefore, his interest as well as other thinker's in the descriptive research program of rationality and decision through the cognitive processes that determine the behavior of the rational decision-making and problem-solving individual (Sternberg, 2000, p. 31). So, the central point that separates Simon's descriptive program on one side from the normative and behaviorist approaches on the other side is related to the affirmation or negation of an empirically sustained psychology of decision, and more specifically to the contention of internal determinants as the cause of the behavior of rational decision-making individuals (Simon: 1976, p. 129).

2. The Theoretical Core of the Descriptive Program of Rationality and Decision

A primordial concern of the Simonian descriptive program of rationality and decision is to provide a proposition that enables to postulate, or even reconstruct, behavior or rational decision. However, from this perspective, the model of the rational individual does not incorporate a priori the capacities that enable him or her to rationally solve the problem of decision as in the normative program. Furthermore, in the Simonian descriptive approach the model of the rational individual internalizes a series of restrictions to rational action which are interpreted in terms of the assumption that his or her cognitive capacities are limited, which is far from imputing irrationality to him or her. Thus, whereas normative thinkers suppose perfect calculation capacities, descriptive thinkers put limits to these capacities, incorporating these limitations into their model of the rational individual. For this reason, the empirically constructed model of the rational individual must be subordinated to certain empirical restrictions, in this case, the restriction which holds that cognitive calculation capacity is a scarce resource.

In this sense, while the rationality of the normative program is theoretical and derives completely or a priori from its model of the rational individual, given that all the conditions to guarantee rationality are found inside the model, in the descriptive model rationality is deprived of its role as a theoretical presupposition. It is instead obtained through the theoretical reconstruction of the notion of rationality from the analysis of empirical evidence. Thus, the theoretical construction of the normative proposition of rationality and decision seeks a solution which is already contained in its model of the rational individual. The theoretical construction of the descriptive program is built with the purpose of solving the problem of how rationality is still possible given the cognitive limitations verified on the basis of empirical evidence and incorporated into its model of the rational individual. For this reason, since the central thesis defended by the descriptive program is the regulating idea that decision is rational, its empirical support must be sought in the reconstruction of the

interaction interface or *process* between the two agents supposed within the core of this program, the rational individual and the choice environment, as well as in the analysis of the decision process considered as a whole and not only focusing on its result, decision. As it is not aprioristic, this version of the notion of rationality defines it as the possibility of adapting means to ends (Simon, 1965).

2.1. Decision Setting: the Model of the Rational Individual and the Choice Environment

Having made clear that the premisses which uphold the model of the rational decision-making individual do not support a priori the conditions to produce optimal choices, that cognitive restrictions besides budgetary ones (or legal or external in any other kind) outline his or her opportunity set and consequently his or her final choice, and moreover that the concept of rationality must be empirically reconstructed, we present next the main components of the decision setting: the model of the rational individual and the possibility of an interface between him or her and the choice environment. Thus, the main unit of analysis of the proposition formulated by Simon is not the isolated rational individual but a system which includes the rational individual and the task environment, more particularly here the choice environment.

In this system two intimately related components are identified: the organism (or rational individual) and the environment (choice environment)ⁱⁱⁱ. Simon ascribes to these components the same properties considered in the general structure of decision theory, *i.e.*, an individual defined according to his or her objectives and an environment organized on the basis of the resources available in it. Those resources are called alternatives in a decision setting and the available set of these alternatives is called an opportunity set. The central thesis regarding the interaction between the two components of the decision setting refers to the relation between them and says that ***the rational individual shows an adaptive behavior, i.e., has the capacity of adapting to the environment.*** It is in their adaptive capacity that ultimately lies their rationality, which is an adaptive rationality. This justifies the definition of rationality as an adequation of means to ends, instead of an optimization of the results as the classical conception of rationality and decision defined it.

Nevertheless, because of the limits imposed upon this capacity and given the complexity of the environment, this adaptation will not create a perfect match but only an approximate one between the rational individual and the choice environment. For this reason, rationality for Simon is also bounded rationality. So, it must be specified that, in the program initiated by Simon, the capacity to behave adaptively, *i.e.*, to provide outputs adequate to the task environment or choice environment, constitutes the true expression of the organism's intelligence or rationality (Simon, 1990, p. 7). Thus, for Simon an intelligent organism is a programmable system (Simon: 1969, 1982). And given that rationality or intelligence is equated with adaptation, Newell and Simon (1972) propose to draw up a specific measuring scale for intelligence. Each scale represents the ability of the individual in a given category of tasks (Newell and Simon, 1972, p. 81-82).

The exposition that follows seeks to explicitly define the adaptive capacity which is being considered. According to Simon (1982), there are three levels of adaptation. The first is introduced by evolutive biology. In this case, it is the adaptation generated by the environment over the species and it operates by means of two forces, mutation and natural selection. The result is the selection of those species *genetically* adapted to survive in a given

environment. Adaptivity is measured here in terms of reproduction rate and perpetuation of the species. Here the adjustment cannot be supposed as deliberately established. The second level of adaptation is called learning and involves control by the individual, although intentionality or conscience are not necessary conditions for learning. In this case, the individual is expected to mold to the environment thanks to the accumulation of information that he or she gathers during his or her life. In this case, it is a long-term adaptation whose reasoning and learning mechanism is induction, which discriminates two steps in the learning process: (i) one step involving the accumulation of information; (ii) another step of recognition of patterns based on that information.

But Simon sees another even more fundamental level of adaptation. In this category are included performance activities, which represent the possibility of short-term adaptation, *i.e.*, implemented *during the execution* of a task. It is this kind of adaptation that Simon is interested in. In the proposition organized by him and Newell, the adaptation undertaken during performance does not exclude learning; on the contrary, it expects that this will be one of its secondary consequences (Newell and Simon, 1972, p. 7). In this case, performance is the main effect and its importance lies in the readiness with which the individual adjusts to the environment. And the execution activity that Simon examines is the task of problem solving, because of the challenges that it raises for the individual's cognitive/adaptive capacity. To understand better what enables this adaptation it is necessary to incorporate into this model of the rational individual a microscopic level that deals with the individual in purely psychological terms, that is to say, according to his or her internal organization. The central thesis of this proposition is formulated by Simon as: "there is an inner environment..." (1969, p. 52).

The assumption of an individual endowed with an internal environment is the mainstay of all cognitive science and provides the foundation that justifies the introduction of a psychology of decision within the descriptive program of rationality and decision. In this level, Simon's program diverges from the behaviorist approach of behavior analysis, even though both feature an aspiration to empirical adequation. To reconstruct the psychology of this internal environment, Simon uses the analogy between the cognitive process of the human mind and the processing of a computer. Thus, the internal environment is procedurally defined by Newell and Simon as a physical symbol system or an information processor. In operational terms, a physical symbol system is described by Simon as follows: "A physical symbol system is a system capable of inputting, outputting, storing, and modifying symbol structures, and of carrying out some of these actions in response to the symbols themselves" (Simon: 1990, p. 3).

These operations of the physical symbol system can be translated in psychological terms as perception (*inputting*), behavior and environmental response (*outputting*), memory (*storing*) and learning (*modifying symbol structures*) capacities. These operations, which define the model of the rational individual in the descriptive program, are made possible by the structure present in this system, which has the following components: (i) memory; (ii) processor; (iii) receptors; (iv) effectors. The computer also has structures and functions similar to those contained in the physical symbol system. Thus, according to his or her operational and structural features, the decision-making individual is constructed as a physical symbol system in analogy with the computer, taken by the descriptive program as an adequate model to provide *insights* and a better understanding of the information processing mechanism (reasoning) of the individual^{IV}.

Besides the characteristics positively attributed to a physical symbol system, a central property assigned to it consists in the presupposition that its operational capacities are limited.

Consequently, unlike the normative approach, the descriptive approach does not deal with an individual with properties ideally adequate to solve problems and make decisions. The hypothesis of the limitation of the operating capacities of the individual corresponds to the aspiration to empirical adequation because, as we stated before, the empirical individual's behavior deviates from the normative axioms and postulates. Still regarding its operability, a physical symbol system depends on two primitive work units: (i) *symbols*; (ii) *elementary information processes* (EIPs).

Symbols are internal representations that originate from the environment or, more properly, that act as inputs which feed the system (Newell and Simon, 1972, p. 7). Hence, the semantic or substantive aspect with which the information processor operates derives from the relation between the internal system of the individual and the external environment where he or she is inserted. The semantic elements or symbols are expressed by means of a language of states and consist in signification units gathered from simple or multiple relations to form symbolic structures. According to their relations, these structures present themselves as a series of organizations, for instance as lists or trees, or according to associations, attribute-value relations being one of those cases. The symbolic structures formed by means of these relations are stored in a semantic memory. An important property of the structure, but not of the symbol, consists in its designatory capacity, *i.e.*, the ability of linking to a referent (Newell and Simon, 1972, p. 24). So, symbols are referents and the designatory capacity is a property of the symbolic structure through which access to a symbol leads to access to other symbols connected to it in the symbolic structure.

The other component of the system, the elementary information processes (EIPs), is made up of operations which act upon the symbols making calculations based on them. They institute, in their turn, the syntax of the system. Thus, while the syntax establishes the physical symbol system as a formal system, the semantic enables the interaction of this system with the external environment. EIPs are expressed in a process language and consist of basic operations that do not suffer further fragmentation. An EIP, unlike a symbol, is not determined by the environment but by the processor's nature, *i.e.*, it has an internal origin (Newell and Simon, 1972, p. 7 and 29).

In their turn, EIPs can be combined to form composite processes. Elementary or composite information processes are called operators since they *operate* with symbols. Some elementary information processes established by Simon are, for instance, to register, copy, move, delete and compare symbols (1969). The combination of a series of processes according to rules creates a program. Programs form a special class of symbolic structure because, in this case, they represent a process and not a state structure. Since they are represented as symbolic structures, it is possible to manipulate them, *so that a program can be constructed or altered*, just as symbol or state structures are constructed and altered, through its interaction with the environment. Moreover, programs are not exclusively constituted of processes as they also have symbols and symbolic structures. Therefore, whereas EIPs are supposed to be a part of the system's nature, *i.e.*, of the individual's internal environment, symbols are internal entities determined by the environment external to the processor. Programs, in their turn, are empirically constructed, uniting symbols and processes through the interaction of the internal system with the environment, forming specific programs for the execution of certain task types. These programs are, consequently, specialist programs or domain-specific programs (Simon, 1976).

Due to this characteristic, domain-specific programs are called *empirical rules* or *heuristic strategies*^v. These rules form a repertory of strategies to solve problems which require

analysis^{vi}. Its central characteristics consist in the selectivity with which they conduct the research within the space of the problem, in the internal representation of the task environment, and in the simplifications adopted to this end. However, although we have spoken here of heuristics instead of algorithms, therefore of empirical rules which are neither rigorous nor deductive and which adopt selectivity to simplify and research information within the space of the problems, generating outputs – the behaviors – which are satisfactory but not optimal, we may assume that, for Simon, they are intelligent rules, since their selectivity seeks to concentrate only on the information relevant to the problem. Besides, the selectivity typical of these programs or heuristics is empirically constructed through the exposition of the individual to certain task environments that guide the construction of these programs.

As to applicability, the programs can be of two types: (i) general-domain; (ii) specific-domain (Simon: 1990). Moreover, as to their specificity, programs are specific in relation to: (i) the domain within which they operate; (ii) the inputs that trigger the program. In this way, the heuristic strategies empirically constructed on the basis of symbols and elementary processes vary as to their generality but are commonly specific-domain, *i.e.*, they apply to circumscribed situations. Specific-domain programs contain processing rules which act selectively on the states and generate outputs adequate to the objective specified within the program. In this level of analysis, this approach incorporates functional explanations, given that a program is organized and justified according to the ends which it serves. Thus, each specific-domain program is supposed to be directed toward an end. Simon supposes that the individual has a portfolio of programs or heuristic strategies of this kind. Because of these characteristics, we can say that, while behaviorists favor efficient causes, that is to say, behavior determinants, supposing that they are always external to the cognitive apparatus of the individual, Simon associates efficient causes to final causes, *i.e.*, the function or purpose of behavior, which guides and organizes the construction of these programs or empirical heuristics. The latter must be executed only if certain environmental conditions are verified.

One notes that the institution of specific-domain programs empirically constructed on the basis of the interaction of the individual with the task environment leads Simon's proposition to frontally oppose the normative program of rationality and decision, according to which the rules contained in the axiom framework are universal or independent from content. Simon says about the programs constructed according to normative models: "...it has been assumed to be describable in a fashion that is completely independent of information about those other decisions or about any other aspect of the real world" (1983, p. 15).

In this last case, the rules are established deductively. On the other hand, in the descriptive program, if the individual thinks and behaves according to a portfolio of empirically constructed programs or heuristic strategies, the structuring of those mechanisms depends on the environment where the decision-maker is inserted, for it is this environment that provides the inputs which feed the physical symbol system. Therefore, what matters is to know whether the system is situated in the Pleistocene, the Middle Ages or the Modern World.

On the basis of this discussion, it may be verified that the individual thinks and behaves through empirically constructed heuristic strategies or programs. Nevertheless, his or her psychic structure is equated to the model of a computer. According to Haugeland (1987) a computer is always a formal system. In other words, the main characteristic of a physical symbol system, which might be a machine or the individual's psychic structure, still maintains the formalism defended by the normative program: "A computer is an interpreted automatic formal system. (...) A formal system is like a game in which tokens are manipulated according to rules, in order to see what configurations can be obtained" (1987, p. 48).

In this sense, Quiggins (1993) affirms that a formal system or, in his wording, an axiom system, does not make any reference to the external environment. For this reason, although the formal system is fed by external inputs that determine the symbols internally, according to Haugeland a symbol is not a formal element of the system. For him “1. The meanings of simple symbols (e.g., words) are arbitrary; and 2. The meanings of complex symbols (e.g., sentences) are systematically determined by their composition” (1987, p. 91).

This property seems incompatible with the notion of empirical rules constructed on the basis of the interaction between the physical symbol system and the environment. Still, Haugeland’s assertion does not differ from Newell and Simon’s (1972) according to which a symbol does not have a designatory capacity, this being solely a property of the symbolic structure, since a symbol does not connect itself to an object of the world through rules but only in an arbitrary way. In other words, given the inputs that feed the system, every ulterior processing is predetermined by its internal rules. Nevertheless, a formal system does not necessarily work with logical rules, which stipulate only universal relations, neglecting any reference to the specific contents that feed the system. The physical symbol system considered in Simon’s (1979, 1990) and Newell and Simon’s (1972) work operates according to conditional rules, which incorporate the syntax and express the admissible relations presupposed by a formal system *but are also sensitive to inputs from the environment*, which correspond to the semantic element. A logical rule of the first kind can be represented by a syllogism, for instance, the Barbara syllogistic form:

EVERY x IS y

EVERY w IS x

EVERY w IS y

In this case, only the relations legitimated by the system are established without any mention to external inputs since these relations are valid for any x , y and w . Conditional rules, with which the physical symbol system laid down up to now works, are constructed by means of: (i) symbols, which determine their conditions; (ii) actions, which are operators applied to the symbols. The structure of these rules can be represented as:

CONDITIONS (*inputs*) → ACTION (*output*)

Thus, action depends on the end sought by the individual (in this sense, the system may be said to be totally determined and closed), but also on the conditions of the task environment (here the system is characterized as an open system). What Simon does here in the explanation of human behavior is nothing but the combination of internal and external determinants or, more precisely, of final and efficient causes. Thus, unlike the behaviorists, who favor efficient causes, the environmental determinants of behavior, Simon seeks to associate these to the internal determinants as well, namely the function or purpose of behavior, which appears as its final cause. In this sense, behavior and empirically constructed heuristic strategies are the resultant of the interaction between the internal environment, which sets the end, and the external environment, which provides the conditions that must be met in the environment so that, once the information processing or behavior is executed, the end is attained.

The distinction between the logical rules model and the empirical rules model is considered by Thagard (1998). According to him, conditional rules are less rigid and therefore suitable to represent and interact with the external environment. Logical rules, on the other hand, are universal relation and operation patterns independent from content or external

environment^{vii}. In other words, the difference between a logical rule and an empirical rule is that the first determines only the relations admitted by the system whereas the second establishes relations between actions *and contents*. Therefore, despite its being treated as a formal system, the physical symbol system continues to be deeply tied to the external environment that surrounds it because its organization depends on this interaction with the environment, given that the programs empirically generated by it to execute specific tasks depend on inputs from the environment (Newell and Simon, 1972, p. 79).

Having provided considerations about the psychology of the individual, *i.e.*, the first entity to be presupposed in the program, we shall deal next with the other component supposed in the decision setting. This component, called task environment, is structured on the basis of an end plus the conditions or environmental characteristics that are relevant to realize or reach this end (Newell and Simon, 1972). Hence, in a chess game, for instance, what matters are the configurations of the board (states) and the admissible movements (processes) to obtain a checkmate (end), but not the material of which the figures are made nor the size of the board. The activity of reaching ends developed in the task environment is called problem-solving activity, and decision-making is one of the categories among problem-solving activities, that of decision problems. A problem is represented internally, within the physical symbol system, in a *problem space*. Once the model of the rational individual and the task environment are characterized, as well as the adaptive form in which the first interacts with the second, the following exposition proceeds to the reconstruction of the main values and ends upheld in the Simonian approach which guide his theoretical construction of rationality and decision.

2.2. Axiological Restrictions to the Theory of Rationality and Decision

Proceeding with the rational reconstruction of the theoretical core of the Simonian approach, we shall deal in this section with the main values and ends which guide and restrict the construction of his descriptive approach of rationality and decision. According to the version of the methodology of the theory of science adopted here, these restrictions in the form of values and ends shall be presented by the means of theses, and these, in their turn, can be of two types: (i) end theses; (ii) value theses. Here, the first point to the objectives pursued by Simon's theoretical construction and the second to the axiological restrictions on which it is erected.

[Values] The possibility of an adequation between the rational individual and the environment, expressed respectively through ends (put by the individual) and means (provided by the environment), is fundamental to the theory of rationality and decision. This adequation consists in the possibility of reconstructing rationality as an empirical concept. Therefore, we can say that decision is defined as a question of means to ends, and rationality as the set of rules (empirical rules here) that enables to justify and intermediate this adequation. On the basis of this initial assertion, the first thesis upheld here is that rationality is a value to be defended within the program. However, its conditions of preservation are not guaranteed a priori by ontology, as it is the case within the normative program, which assumed a rational individual capable of deliberating properly, *i.e.*, with full capacity to do so.

Having posited rationality as a regulating idea, it must be shaped according to the evidence obtained from the analysis of how the individual really makes decisions, according to the Simonian program's aspiration to empirical adequation. This evidence consists in restrictions incorporated into the model of the rational individual and can be expressed, once the individual is shaped as a processor of physical symbols, in terms of the assertion that his

or her computational capacities are limited. Having admitted limited computational capacities, *i.e.*, cognitive capacities as a scarce resource, the process involved in the solution of a problem, be it a decision process or not, entails cognitive costs incurred in this process, costs which, for a processor with unlimited cognitive capacities, would not need to be incorporated into the model of the rational individual, as it is the case with the model constructed in the classical conception. These procedural costs are characterized in terms of cognitive effort. Thus, costs related to the processes derive from the cognitive effort exerted in the execution of tasks involving superior cognitive capacities, as it is the case of the activity of solving problems and making decisions.

Taking these restrictions into account, it can be verified that, given its empirical aspiration, the rationality established by the descriptive program must undergo some reformulations compared to that of the normative program. Consequently, both the normative and the descriptive programs uphold rationality as a value, although in the normative program it is treated as if it were a piece of data, thus belonging a priori to the model of the rational individual. In the Simonian approach, however, rationality is empirically reconstructed; it is not, consequently, a piece of data or an a priori assumption of the theory. In an ample sense, if it can be imputed to both programs of rationality and decision, the normative as well as the positive, rationality may be understood as the set of rules that makes it possible to justify the adequation between means and ends. However, the nature of those rules differs in the two programs. In the normative approach, adequation between means and ends is demonstrated by the theory according to logical rules, the postulates and axioms. These rules, together with the optimization principle, enable to postulate the adequation between means and ends in optimal terms, given the values, preferences and weights/beliefs (inputs) provided to the alternatives by a given decision-maker. Because of this nature, it is assumed that the decision problem only has one solution, or more precisely, the selection of only one alternative can be justified – that with the highest utility or expected utility.

According to the descriptive program, however, the result of the adjustment between means and ends, according to the interpretation of rationality as bounded and adaptive, cannot be logically deduced from theory. On the contrary, rationality is defined on the basis of the capacity of the decision-maker to adapt to the environment. In other words, since the decision-maker is not endowed a priori with an apparatus that guarantees an optimal adequation between means and ends, the adequation between means and ends is empirically reconstructed, as it cannot be deduced necessarily from theory. The descriptive program of rationality and decision does not renounce, however, to the regulating idea of rationality as the possibility of adjusting means to ends. But this adequation derives from the adaptive capacity of the decision-making individual. Nevertheless, due to the restrictions mentioned, adaptation results not in perfect but only approximate integration between means and ends. To quote Simon “Because of the limits on their computing speeds and power, intelligent systems must use approximate methods to handle most tasks. Their rationality is bounded” (1990, p.6).

In accordance with this proposition, the relation between means and ends is established weakly, for, in this case, it is empirically reconstructed instead of demonstrated by means of postulates and axioms^{viii}. According to Simon (1990), the mechanism that enables this reconstruction and the postulation of rationality, although bounded, consists in its qualification as adaptive rationality, which is responsible for the possibility of interaction between the rational decision-making individual and the choice environment. In this way, because of the cognitive or computational restrictions attributed to the decision-maker, adaptation does not operate in optimal terms, but its results are only approximative in relation

to those obtained by the optimization principle. This because the latter did not take into account in their calculation techniques the costs related to the processing of the information necessary to decision, that is to say, the cognitive effort involved in the process of decision-making, but only the costs related to the alternatives, for instance the price of the alternatives or the cost of the opportunity of choosing one alternative instead of others. It is for this reason that Simon replaces the postulate of perfect rationality by that of bounded rationality.

On the other hand, according to Simon, if the costs related to the processes involved in decision are computed together with the costs of the alternative, that is to say, if the optimization principle is more widely applied, considering *the processes and the product*, this principle could be recovered:

To be sure, we can formally view these as optimizing procedures by introducing, for example, a cost of computation and a marginal return from computation, and using these quantities to compute the optimal stopping-point for the computation (1976, p. 140).

For Bordley (1985), the total utility involved in the decision setting should incorporate the utility arising from the alternate choice according to a certain strategy and deduct the costs involved in the execution of the strategy (computational or procedural costs or, in psychological language, cognitive effort). Consequently, in this proposition process costs would be taken into account in the utility calculations and it would not be necessary to abandon the optimization principle, only to redimension it. Furthermore, in the descriptive program the optimization principle applied to the result (having deducted procedural costs) is followed to the letter only in the cases involving well-structured problems. In this sense, it draws nearer to the normative program. However, for the remaining problems, Simon says that: "...this procedure may represent a sufficient approach to optimization..." (1979, p. 12).

The greater theoretical problems generated within the program, which intends to guarantee support to the concept of rationality, arise from its aspiration to empirical adequation. This ambition is asserted in a second value thesis. The search for empirical adequation commands the theoretical construction so that the theory must adapt to the empirical evidence found. The theoretical problem which then arises is how to weave this adequation between theory and empirical evidence. This proposition differs from the normative perspective, according to which problems are solved formally and a priori by means of the instruments established by the theory. Here the goal is to follow the criterion of correspondence between what is established in the program and the studied phenomenon without abandoning rationality as a regulating idea, what separates, as we said, Simon from other descriptive propositions of decision. Hammond discriminates the criteria of coherence and correspondence as patterns that enable normative and descriptive thinkers to assess the decisions:

The goal of a correspondence metatheory is to describe and explain the process by which a person's judgments achieve empirical accuracy. The goal of a coherence metatheory of judgment, in contrast, is to describe and explain the process by which a person's judgments achieve logical, or mathematical, or statistical rationality (2000, p. 53).

These same criteria used to assess decisions can also be assigned to theory, so that while normative thinkers favor the formal structure of the models proposed by the theory, descriptive thinkers are concerned with their empirical adequation as well. Thus, the maximum value of decision theories, *i.e.*, rationality, is sustained by the descriptive program under the condition that its preservation does not impair the criterion of correspondence or empirical adequation. In other words, the theoretical construction derives from a commitment

or tradeoff between the preservation of the regulating idea of rationality and the aspiration to empirical adequation, and the idea of the individual's adaptivity appears as an intermediary element in this theory. The individual's adaptivity is operationalized in the program's methodology through the analogy between the *modus operandi* of the decision-makers and the functioning of an information processor. The rational individual is identified, in this sense, with a physical symbol processor, as we said before.

[Ends] It is because of those concerns about the values sustained by the descriptive program of rationality and decision – rationality and empirical adequation – that we affirmed in a first end thesis that the objective of the program is to construct a psychological theory of the decision-making individual. This theory is called theory of procedural and adaptive rationality (Simon: 1976, 1988, 1997). The theory of procedural and adaptive rationality includes the institution of a model of the rational individual according to the criterion of correspondence with the empirical data, just as the behaviorist proposition does. But, unlike the latter, it incorporates internal determinants into the explanation of the behavior adopted by the decision-maker. This means that, in opposition to the behaviorist school, for Simon's descriptive program the determinants of behavior are not limited to environmental determinants but also refer to the specific properties of the decision-maker, particularly those related to his or her cognitive capacity and to his or her manner of processing information (hardware and software), adopting the computer as an analogy to define his or her structure and cognitive functioning as well.

The reconstruction of the internal processes should also make possible the postulation of rationality by means of the supposed adaptive capacity attributed to the individual. This capacity enables to assume that alternative information-processing resources incorporated into the model of the rational individual can be made available and used by him or her according to the demands of the task environment. These alternative processing resources are heuristic strategies (rules of thumb) attributed to the individual who, assuming that he or she has a portfolio of heuristic strategies, tackles the demands of the problem accessing the strategy that seems to be more adequate to the case^{ix}. There lies his or her adaptive capacity: in accessing the heuristic strategy that is more adequate to the demands of the problem. Moreover, the heuristic strategies incorporated into the descriptive program of rationality and decision are reconstructed through the observation of how experimental individuals solve problems and make decisions in laboratory situations. And above all, the method of information retrieval adopted consists of the protocols, which are a methodological resource fairly common in psychology, abandoned by the behaviorists but recovered by Simon.

Given the need to reinforce the notion of the adaptive capacity attributed to the individual, the construction of this taxonomic framework of heuristic strategies consists in one of the ends that are sought by the theories of the descriptive research program of rationality and decision (Simon, 1990, p. 6-7). Parallely to the objectives of establishing a psychology of the decision-making individual, called theory of procedural and adaptive rationality, and of creating a taxonomy of the heuristic strategies that embody the qualification of the adaptivity attributed to the individual, we can establish as a third end thesis that the goal of the theory is also to construct a theory about the demands of the task imposed by the environment. These demands are treated in the theory of substantive rationality, according to Simon (1976, 1988, 1997). Therefore, while the theory of procedural rationality is intimately related to the processes and mechanisms attributed to the rational decision-making individual, the theory of substantive rationality must define the structure of a series of problems that arise, setting their invariables and specific contents. Because of this, the theory of substantive rationality must provide a taxonomy for a large set of problem-

solving activities^x. Among those activities is the class of decision problems and its subcategories. Given that the descriptive program considers that the demands which characterize the task or choice environment are fundamental in the selection of the heuristic strategy that is more adequate to the execution of the activity, these two theories are deeply interwoven. It is also for this reason that it makes sense to establish different scales of intelligence that measure the performance of the decision-maker in different classes of problems.

Conclusion

The main thesis defended in this paper affirms that in spite of its attempt to construct an empirically based individual decision-maker, the descriptive program of rationality and decision launched by Simon does not abandon the notion of rationality as a regulating idea. Because of this determination, it interprets rationality as bounded and adaptive. Bounded because the empirical adequation of its theoretical individual to the empirical individual requires the incorporation of the thesis that the latter has limited cognitive or computational capacities, these being its main scarce resources. Adaptive because, although these capacities are limited, the individual is capable of responding contingently to the task environment, guaranteeing the adequation between means and ends, which is the very definition of the notion of rationality. It must be highlighted, nonetheless, that this adequation occurs only in an approximate way because procedural costs (cognitive costs) are then taken into account in the decision-making process.

References

- BELL, RAIFFA and TVERSKY. "Descriptive, Normative, and Prescriptive Interactions in Decision Making". In: D.E. Bell, H. Raiffa, A. Tversky (eds.). *Decision-Making: descriptive, normative, and prescriptive interactions*. Cambridge: Cambridge University Press, p. 9-30, 1988.
- CHIAPPIN, J.R.N. "Racionalidade, Decisão, Solução de Problemas e o Programa Racionalista". *Ciência e Filosofia*, nº 5, p. 155-219, 1996.
- BORDLEY, R.F. "Systems simulation comparing different decision rules". *Behavioral Science*, vol. 30, no. 4, p. 230-239, 1985.
- HAMMOND, K.R. "Coherence and Correspondence Theories in Judgment and Decision-Making". In: T. Connolly, H.R. Arkes and K.R. Hammond (ed.). *Judgment and Decision-Making: an interdisciplinary reader*. Cambridge: Cambridge University Press, p. 53-65, 2000.
- HAUGELAND, J. *Artificial Intelligence: The Very Idea*. Cambridge: MIT, 1987.
- KAHNEMAN, D. and TVERSKY, A. "Prospect Theory: an analysis of decision under risk". *Econometrica*, vol. 47, p. 263-291, 1979.
- LAKATOS, I. "O Falseamento e a Metodologia dos Programas de Pesquisa Científica". In: I. Lakatos and A. Musgrave (orgs.). *A Crítica e o Desenvolvimento do Conhecimento*. São Paulo: Cultrix/EDUSP, p. 109-243, 1979.

- LEISTER, C. *Programa de Pesquisa da Racionalidade e Teoria da Decisão: Racionalidade e Teoria da Decisão Adaptativa*. Dissertação de Mestrado. FFLCH-USP, 2001.
- NEWELL, A. and SIMON, H.A. *Human Problem Solving*. Englewood Cliffs: NJ., Prentice Hall, 1972.
- PAYNE, J.W., BETTMAN, J.R., and JOHNSON, E.J. *The adaptive decision-maker*. Cambridge: Cambridge University Press, 1993.
- POZO, J.I. *Teorias Cognitivas da Aprendizagem*. Porto Alegre: Artes Médicas, 1998.
- QUIGGINS, J. *Generalized Expected Utility Theory*. Boston: Kluwer Academic Publishers, 1993.
- SIMON, H.A. *Administrative Behavior*. New York: The Macmillan Company, 1965.
- _____. *The Sciences of the Artificial*. MA: MIT Press, 1969.
- _____. *From substantive to procedural rationality*. In: S.J. Latsis. *Method and Appraisal in Economics*. Cambridge: Cambridge University Press, p. 129-148, 1976.
- _____. *Models of Thought*. New Haven: Yale University Press, 1979.
- _____. "Cognitive Science: The Newest Science of the Artificial". In.: D. Norman (org.). *Perspectives on cognitive science*. New Jersey: ALEX Publishing Corporation, Norwood, 1982.
- _____. *Reason in Human Affairs*. Stanford: Stanford University Press, 1983.
- _____. "Rationality as Process and as Product of Thought". In: D.E. Bell, H. Raiffa, A. Tversky (eds.). *Decision-Making: descriptive, normative, and prescriptive interactions*. Cambridge: Cambridge University Press, p. 58-77, 1988.
- _____. "Invariants of Human Behavior". *Annual Review Psychology*, vol. 41, p. 1-19, 1990.
- _____. *An Empirically Based Microeconomics*. Cambridge: Cambridge University Press, 1997.
- SIMON, H.A. and HAYES, J.R. The Understanding Process: Problem Isomorphs. *Cognitive Psychology*, vol. 8, p. 165-190, 1976.
- SKINNER, B. F. *Ciência e Comportamento Humano*. São Paulo: Martins Fontes, 1989.
- STERNBERG, R.J. *Psicologia Cognitiva*. Porto Alegre: Artes Médicas, 2000.
- THAGARD, P. *Mente: Introdução à Ciência Cognitiva*. Porto Alegre: Artes Médicas, 1998.

NOTES

ⁱ In the 19th and 20th centuries, with Duhem's dynamic realism, the logical positivism of the Viennese Circle and Popperian falsifiability, scientific methodology separated the fields of the method of discovery, which proposed methodological approaches to the construction of theories, and the method of justification, concerned only with the testing of theories, and kept within its reach only the latter, relegating the former to the study of psychology. Simon is identified here as one of the thinkers who recovers the field of the methodology of discovery and brings it to the center of the scientific debate.

ⁱⁱ The outlining and differentiation of the core and heuristic of these two programs were done with greater detail by Leister (2001).

ⁱⁱⁱ The fragmentation is done for didactical ends since the theory defines these two components exclusively on the basis of the relation that they have between them.

^{iv} This analogy seeks to make clear the similar functional properties of the individual and the computer, considering the computer's software to explain the *modus operandi* of the mind. However, the same does not apply for the hardware, since the physical structure of the computer cannot be taken as analogous to that of the human being, namely its brain. Computational models based on neural networks (connectionist models), on the other hand, have been making an effort in this direction. But Simon believes that psychology and neurophysiology work with distinct levels of explanation that are not interchangeable:

We do not believe that this functional equivalence between brains and computers implies any structural equivalence at a more minute anatomical level (e.g., equivalence of neurons with circuits). Discovering what neural mechanisms realize these information-processing functions in the human brain is a task for another level of theory construction. Our theory is a theory of the information process involved in problem solving, and not a theory of neural or electronic mechanisms for information processing (1958, p. 163).

^v Heuristics are the methods derived from inductive reasoning that allow to make generalizations from observed events and to discriminate their covariations (Sternberg: 2000). As they are inductively constructed they also include symbols and contents. Their power lies in the possibility of constructing specialist systems.

^{vi} According to Simon (1990), problem solving includes a series of mechanisms separated in three categories, namely: (i) recognition processes; (ii) heuristic research; (iii) recognition of serial patterns. In the first case problems are solved without the need of analysis, on the sole basis of the knowledge stored in memory. In the two other classes an analysis of the problem is required. In heuristics, it occurs by means of selective research. The recognition of patterns includes both the retrieval of knowledge stored in memory and analysis.

^{vii} In this sense, Thagard affirms that "...the people who develop systems based on rules were happy to lose a little of the rigor of logic-based systems, thanks to the increase in computational power. An advantage comes from the fact that the rules do not have to be interpreted as universal truths" (1998, p. 50).

In this case, the increase in computational power provided by empirical rules refers to the property that, being a specific-domain program, its answers become more efficient thanks to the sensibility in relation to the particular elements of the problem.

^{viii} One of the consequences of this weak relation between means and ends is the impossibility of the theory to predict the result of the decision, since the selected alternative cannot be logically or necessarily justified based on the theory.

^{ix} We have, therefore, two levels of decision: (i) the decision related to the alternative taking into account the opportunity set; (ii) the decision about the heuristic strategy that is more adapted to the choice environment which the decision-maker faces. The selection of the heuristic strategy, as we shall see in a subsequent paper, includes a metarule that defines the optimal balance between process costs and result benefits. As said before, if applied to this second level of decision, *i.e.*, if the process costs were considered, the optimization principle could be recovered in the decision theory while keeping the empirical adequation of its decision models.

^x The concern to establish problem classes led Simon to work with isomorphic problems, which delimitate a problem category mainly by following their structural similarities (Simon and Hayes: 1976).