PSL Quarterly Review, vol. 70 n. 279 (March, 2017), 7-33

Paths in contemporary economics and sciences of artificial that originate from Simon's *bounded rationality* approach

MASSIMO EGIDI*

1. The origins of the "bounded rationality" approach

Simon first developed the idea of bounded rationality in his PhD dissertation,[1] later published in the book *Administrative Behavior* in 1947. Ironically, the strength of this book originated from the author's lack of experience in the field of organizational sciences. Given his lack of familiarity with the practical aspects of organizational problems, Simon was inclined to investigate more abstract and general aspects of human behavior within organizations. He did so, at first, by analyzing organizations as goal-oriented structures characterized by the division of goals and the coordination of tasks. This context was a good starting point for studying the limits of human decision-making, even though it was clear to him that these limits are a general characteristic of human decisions:

> "[t]he idea of bounded rationality, which appears to be the most novel and original component of the work, is not specifically an organizational concept. It applies as fully to individual decision making as to organizational decision making. By the age of twenty-five, I had already had ample experiences in life to understand the limits of the economists' framework of maximizing subjective expected utility as applied to actual human behavior. The scantiness of my experiences with organizations posed no particular limit to my development of an alternative approach to decision making." (Simon, 1996, p. 87).

* LUISS Guido Carli, email: megidi@luiss.it.

[1] "The dissertation contains the foundation and much of the superstructure of the theory of bounded rationality that has been my lodestar for nearly fifty years." (Simon, 1996, p. 86).

His approach was supported by important intellectual observations from Chester Barnard's Functions of the Executive (1938). Barnard was both a great expert of organizational problems and an original thinker. Simon explicitly acknowledges his intellectual contribution.

> "Although Barnard did not construct a systematic theory of decision making, much of his discussion was directed at the executive's decision-making processes. […] The other central idea in my book that appears in only muted form in Barnard's is *bounded rationality*. The closest parallel is Barnard's notion of opportunism and strategic factors, ideas that he derives from John R. Commons. Since I had also read Commons, the latter's Institutional Economics may have been a common source for these various conceptions of rationality that deviate from the economists' maximization of subjective expected utility." (Simon, 1996, p. 87, italics added).

When Administrative Behavior was first published, organizational sciences were a field of inquiry entirely distinct from microeconomics, and studies of organizational decisions never referred to the theory of rational decision-making. The firm, in fact, was viewed as a production function, the entrepreneur as a profit maximizer, and the internal characteristics of the organization as a question of optimal planning. The normative role of these optimization techniques made the question of how real managerial decisions take place irrelevant. Administrative Behavior opened up the 'black box' of the internal mechanisms of organizations. Simon identified the main characteristics of managerial decision-making by analyzing the structure of the organizational process. There are many relevant new ideas in the book related to the definition of an organization as patterns of coordination: authority, communication, efficiency, and organizational loyalty. However, the division of tasks and their coordination are more intimately connected with rationality, and thus worth mentioning.

Simon recognized that the core of every organization is the pattern underlying the division of tasks and their coordination: the organization is considered a goal-oriented structure based on internal tasks that must be coordinated to achieve the organization's overall objectives. Behavior within organizations is oriented toward goals, and goals are generally complex and hierarchical. Many intermediate sub-goals must be realized in a specific order for the final goal to be achieved. Additionally, the dynamics of organizational decisions are very complex and have two relevant aspects. First, goals are often defined in very general and ambiguous ways, thus necessitating revision of the sub-goals' hierarchy. Second, many hidden conflicting objectives can be discovered during organizational decisions, when, again, a revision of the sub-goals and their hierarchy may become necessary. It is easy to see that, with this analytical setup, the classical theme of division of labor and coordination would dominate the scenario, leaving utility theory behind.

2. The extraordinary fifties at the Graduate School of Industrial Administration

Administrative Behavior was a 'manifesto' containing new challenging ideas, so, of course, its validation required the possibility to carry out long-term research. The right environment for progressing the new analytical approach soon came with a new initiative in the industrial world. In 1948, the Carnegie Institute of Technology received a donation of $5 million in endowments and $1 million for the building of a new Graduate School of Industrial Administration (GSIA) that would provide business education to students with undergraduate degrees in science and engineering.

> "The donor was William Larimer Mellon, who had founded the Gulf Oil Company. From his industrial experience, he had concluded that modern high-tech firms needed top executives who both were skilled in management and understood science and technology." (Simon, 1996, p. 136).

Soon, Simon was invited to join this venture as a professor of administration and was nominated Chairman of the Department of Industrial Management. The visionary idea of mixing economics, organizational sciences, engineering, and computer science seemed achievable in this new intellectual context.

> "The excitement of the time can be conveyed or re-evoked for those of us who lived through it by listing the labels for constellations of ideas that were born then: operations research and management science, the theory of games, information theory, feedback theory, servomechanisms, control theory (these and others collected under the banner of cybernetics), statistical decision theory, and the stored-program digital computer. […] The ideas were all closely intertwined, with decision making at their core, and they quickly generated a scientific culture, an interlocking network of scientists with a real sense of community, which was almost independent of the special area in which each worked, and which ignored the diversity of their backgrounds and training. They came from physics, statistics, economics, biology, mathematics, engineering, philosophy, and even a few from psychology and political science." (Simon, 1996, p. 107).

Perhaps this approach was overly precocious. Only in recent years have some among the best business schools employed this formula, sometimes due to external suggestions from managers of large innovative companies and the business world. However, during the fifties, the traditional academic fields were distinctly separated, and the new emerging research fields, such as computer science, were considered branches of the old ones (engineering and mathematics), disregarding their pursuit of new goals and the introduction of new methods.

In this context, another academic area was rapidly growing: the theory of decision-making under risk. The theory flourished in the fifties, thanks to the contributions of Friedman and Savage, and was extended to many other areas, such as optimization in linear and dynamic programming. The attention of the great majority of economists was focused on these advancements, which reinforced the central position of the neo-classical approach in economics. As we will see, the theory of decision-making as constrained optimization and Simon's theory of problem solving grew further apart, and it took fifty years, with the birth of behavioral economics, to start an opposite process of critical comparison and intellectual remixing between the two approaches.

As mentioned, the attention of the great majority of economists was focused on the advancements in expected utility theory and, more generally, in the increasing use of formal mathematical models to represent decision-making. Economists were dazzled by mathematics and formal methods, falling for the illusion of gaining the same prestige as the natural sciences just because of the use of formal models. As the history of natural sciences shows, this condition is important (and sometimes necessary), but not sufficient. The condition for sufficiency is, of course, empirical validation, and the road of science is laden with cadavers of models that failed to survive experimental proofs.

Unfortunately, the path taken by mainstream economics in the fifties privileged axiomatization over experimental verification, and only half a century later was the second pillar seriously considered. By ambiguously attributing a positive/normative character to the expected utility theory, Friedman and the Chicago School, of which the former was a leading exponent, exempted scholars from putting their theories under experimental scrutiny. His approach was based on three pillars. First, Friedman suggested an evolutionary defense of 'full rationality' by assuming that those who failed to conform to rational behavior would be gradually excluded by market selection. Second, he claimed that although individuals do not possess the formal tools for calculating optimal solutions, they behave *as if* they do—like billiard players who accomplish complex trajectories with their billiard balls while ignorant of the laws of mechanics. Third, he claimed that all assumptions about individual preferences were irrelevant for proving the theory's validity (Friedman, 1953).[2] This view was widely accepted, and the first important proof of an important shortcoming of the expected utility theory, clearly described by Maurice Allais in 1952, was ignored.[3] Economists avoided this attack on their certainties by re-baptizing this failure a "paradox", thus skipping the problem without solving it.

---

[2] This view opened a heated debate, and was harshly criticized, among many other authors, by Simon (1963).

[3] In 1952, at a symposium held in Paris, Allais presented two studies in which he criticized the descriptive and predictive power of the 'American school's' choice theory, and especially Friedman's version of it (Allais, 1952).

As is well known, the growing continuation of experiments along the lines Allais initiated, along with the development of the theory of bounded rationality, paved the way for the emergence of behavioral economics, thus re-establishing the importance of experimental proof in accordance with the standard methodology of the natural sciences. This process was extremely slow, and took fifty years to accomplish.

During the fifties, the Chicago view of decision-making and optimization increasingly dominated the field, and the bounded rationality approach was marginalized. An increasingly evident conceptual clash was unfolding, with a relevant impact on academic relations.[4] This contrast was particularly apparent within the Graduate School of Industrial Administration (GSIA), which, as a business school, was oriented toward closely connecting concrete industrial management problems with the theoretical models. According to Simon, keeping the balance between the scientific and the professional, between the economic and the behavioral, was an arduous task. For a while, despite continuous contrasting opinions also on a theoretical level, the situation was kept in check. Only the leadership of the chairman of the Economics Department, Lee Bach, allowed to maintain the venture's course and kept the crisis under control.

Simon notes:

> "[t]he problems that created the crisis did not wholly go away; they were built into the fabric of the GSIA mission. One problem was the fascination of abstract mathematical techniques, which sometimes emphasized the mathematics more than the management applications. A second problem was the partial mismatch between the 'pure science' values economists acquire from their discipline and the interest in real-world applications that characterizes a business school. A third, related problem was the near incompatibility of the behavioral theories of economic decision making that some of us were developing with the neoclassical theories espoused by most of the economists." (Simon, 1996, p. 147).

---

[4] "At any rate, I heckled the GSIA economists about their ridiculous assumptions of human omniscience, and they increasingly viewed me as the main obstacle to building 'real' economics in the school." (Simon, 1996, p. 144).

3. The birth of an empirically based theory of decision making in ill-defined conditions

Simon dedicated his first years at GSIA to experimentally validating and further developing the central ideas of *Administrative Behavior*. Because one of the book's central assumptions was to consider organizations as devices coordinating hierarchies of different tasks, it was crucial to experimentally detect the human ability to achieve these tasks.

In 1956, Cyert, Simon and Trow carried out an empirical analysis of managerial decisions that revealed an evident dualism of behavior based on routinized decisions on the one hand, and behavior requiring problem solving and research on the other. This gave rise to a dichotomy between types of decisions, which Cyert, Simon and Trow sharply distinguished as programmed and non-programmed:

> "[d]ecisions in organizations vary widely with respect to the extent to which the decision-making process is programmed. At one extreme, we have repetitive, well-defined problems (e.g., quality control or production lot-size problems) involving tangible considerations, to which the economic models that call for finding the best among a set of pre-established alternatives can be applied rather literally. In contrast to these highly programmed and usually rather detailed decisions are problems of non-repetitive sort, often involving basic long-range questions about the whole strategy of the firm or some part of it, arising initially in a highly unstructured form and requiring a great deal of the kinds of search processes listed above." (1956, p. 238).

From this point of view, the core of the decision-making process is the act of searching and learning, through which individuals acquire the information and knowledge they require to achieve their goals. Simon and his colleagues emphasize that 'rational' decisions are at the foundation of organizational routines, while the essence of decisions is the solution of new problems, along with the connected acquisition of new competencies and skills. Put in a different way, they maintain that problem solving is the fundamental micro-innovative activity that characterizes organizational life—a sort of micro-schumpeterian view.

Of course, under these conditions, the standard choice theory cannot be applied, because the preference orderings are very incomplete—decisions are simultaneously inconsistent and choices are largely ineffective with respect to the goals that are being pursued. The most important part of the process is driven by the subject's ability to formulate and solve problems. This description proved to be very general, paving the way for a formal construction of problem solving theory. Simon went on to develop the formal mathematical tools for representing problem solving activity and, by so doing, sowed the seeds for the development of the field of artificial intelligence.

In 1955, he began his collaboration with Allen Newell, one of the most influential figures among the founding fathers of artificial intelligence. The collaboration gave rise to the creation of new mathematical tools for modeling human problem solving and discovery processes.

> "The most important years of my life as a scientist were 1955 and 1956 [...]. During the preceding twenty years, my principal research had dealt with organizations and how the people who manage them make decisions. My empirical work had carried me into real-world organizations to observe them and occasionally to carry out experiments on them. My theorizing used ordinary language or the sorts of mathematics then commonly employed in economics. Although I was somewhat interdisciplinary in outlook, I still fit rather comfortably the label of political scientist or economist and was generally regarded as one or both of these.
> All of this changed radically in the last months of 1955. While I did not immediately drop all of my concerns with administration and economics, the focus of my attention and efforts turned sharply to the psychology of human problem solving, specifically, to discovering the symbolic processes that people use in thinking. Henceforth, I studied these processes in the psychological laboratory and wrote my theories in the peculiar formal languages that are used to program computers. Soon I was transformed professionally into a cognitive psychologist and computer scientist, almost abandoning my earlier professional identity." (Simon, 1996, p. 189).

As a basis, they shared the project of creating computer programs to solve problems, transferring the principles that guide *human* problem solving to the machine. Newell had attended several courses by George Polya, who, in *How to Solve It* (1945), described problem solving as an activity that transforms the symbolic representation of a problem into its solution. To proceed with this key idea, Simon and Newell first needed to use a computer as a symbol processor.

> "We seized the opportunity we saw to use the computer as a general processor for symbols (hence for thoughts) rather than just a speedy engine for arithmetic. By the end of 1955, we had invented list-processing languages for programming computers and had used them to create the Logic Theorist, the first computer program that solved non-numerical problems by selective search. It is for these two achievements that we are commonly adjudged to be the parents of artificial intelligence. […]
> With that, we opened the way to automating a wide range of tasks that had previously required human intelligence, and we provided a new method, computer simulation, for studying thought. We also acquired considerable notoriety, and attracted critics who knew in their hearts that machines could not think and wanted to warn the world against our pretensions. […][5]
>
> The initial approach also established the precedent, followed in all of our subsequent work, that artificial intelligence was to borrow from psychology, and psychology from artificial intelligence. Thus, Al's programmatic description of his chess-learning proposal, like my 1952 sketch […], used aspiration values and notions of 'satisfactory solution' in evaluating chess moves and discussed the necessity for 'rules of thumb' (later called heuristics) to reduce the enormous search space—the space containing all the branches in the tree of possible chess moves and replies—to a manageable size. By using heuristics and settling for satisfactory moves, the search space could be explored very selectively,

---

[5] "During the summer and autumn of 1957, we gradually converged to a program embodying the newly discovered means-ends analysis. Because the reasoning processes in the program were independent of the particular topic on which it was reasoning, we christened it the General Problem Solver. The general flow diagram of GPS was produced before the end of October 1957, and the planning method (a scheme for simplifying search by abstracting the problem) was sketched a few days later […]. Thirty years of subsequent research has confirmed that means-ends analysis, as embodied in GPS, is a key component of human problem-solving skill." (Simon, 1996, p. 220).

> avoiding any attempt at an impossible exhaustive search." (Simon, 1996, pp. 189-202).

Simon and Newell started a long series of studies on the game of chess, developing programs for a chess computer. This paved the way for studies on understanding the psychology of problem solving. Experiments on chess players and the creation of a theory of problem solving proceeded together, nurturing each other, in the seventies and later. The notion of heuristics was developed, in this context, as the micro-strategies (or rules of thumb) that allow the player (human or artificial) to select the next moves. Their collaboration gave rise to great progress both in the creation of computer simulations of heuristic problem solving and the exploration of heuristics through experiments on the psychology of problem solving. In 1958, they decided to write a treatise on human problem solving based on their research with human subjects and computer simulations. The published volume appeared in 1972, titled *Human Problem Solving.* The book forms a bridge between computation, artificial intelligence, and cognitive psychology.

4. Paths in organizational theory and artificial intelligence

The vast ramifications of the different research paths generated by the bounded rationality approach expanded at different speeds. While some areas expanded at exponential rates, sometimes with disruptive effects, others, due to intrinsic intellectual difficulty or to a smaller number of impacted disciplines, expanded more slowly. While I will privilege the ones that have already caused disruptive effects on the history of economic ideas, I will start by referring to the issue at the basis of Simon and Newell's inspiration, which had slowly been progressing in the literature—the conceptualization of problem solving as changes of representation.

> "During the first six years of my research at GSIA, I filled out, empirically and theoretically, the decision-making framework I had laid down in *Administrative Behavior*. My files yield a planning document that Harold

> Guetzkow and I wrote on February 28, 1952 […]. The most interesting substantive recommendation in the document was that decision making in organizations should be related to learning theory: 'Our work has led us to the conclusion that there is an intimate connection between organizational structure and the learning of frames of reference and roles by members of organizations.' […] The idea foreshadowed the critical importance of reference frames we would now call 'problem representation' or the 'frame problem' in problem solving and learning. Problem representation is still high on the agenda of research in cognitive science today, thirty-nine years after the date of that memorandum" (Simon, 1996, p. 161).

The importance of analyzing changes of representation is evident in, for example, the framing effect theory (Tversky and Kahneman, 1986). Here, individuals change their attitude toward risk based on different representations of the same problem.

In *The Sciences of the Artificial*, Simon discusses representation in the context of design. According to Polya, Simon suggests that all problem solving may be viewed as changes of representation—solving a problem can be conceptualized as representing it so that the solution becomes obvious. He also proposes the development of a taxonomy of representations as the first step toward a theory of representation.

The representation problem has two main aspects, both generated by the General Problem Solver model: (1) the reduction and simplification of the problem space, and (2) its decomposition and modularization.

Perhaps the best example of changes in representation is the missionaries and cannibals problem. Amarel (1968) analyzes the problem by performing a series of about nine different changes of representation to arrive at a problem space where the solution is trivial.[6] He also comments on the possible mechanization of these transformations. On the second aspect of problem representation, i.e. decomposition and modularization, some recent progress seems to indicate promising avenues for future research. Simon's observation that near decomposability is a general property of complex systems has been corroborated by studies on the evolutionary properties of

[6] See also Knoblock (1990; 1991), and Korf (1980).

nearly decomposable organizations and organisms. As Marengo and his colleagues show, near decomposability gives an evolutionary advantage because it allows quick location and ascension to a good fitness peak (Frenken et al., 1999; Simon, 2002; Marengo et al., 2005). However, near decomposability generally entails sub optimality, since decompositions almost inevitably separate into different modules of interacting elements that should be kept together. Thus, the evolution of organizations strikes a balance between adaptation speed and search optimality.

I cannot avoid mentioning a research path that has flourished very much, which is related to organizational decision-making and organizational learning. *Organizations*, co-authored with James G. March and published in 1958, is a synthesis of all the major issues developed in the fifties. The book progressed from the notion of problem solving as individual activity to the notion of organizational problem solving, clearly recognizing organizational adaptation and organizational learning processes within corporations. Identification of these processes proceeded in parallel with the interpretation of division of labor as a problem solving activity, as well as the identification of the recursive division of problems into sub-problems as a property of both organizations and computer programs. The notion of organizational routines and procedures within firms, and their evolution, was born, paving the way for a wave of studies on organizational learning and organizational routines. Further steps in this direction pertained to the development of the notion of organizational routines (Levitt and March, 1988; Levinthal and March, 1993; Cohen et al., 1996; Becker et al., 2005; Feldman and Pentland, 2003), the interpretation of their role in an evolutionary context (Nelson and Winter, 1982), experiments on their emergence from cooperation (Cohen and Bacdayan, 1994; Egidi and Narduzzo, 1997), and the study of their dynamic and evolutionary properties (Dosi et al., 2001; Zollo and Winter, 2002).

By the end of the fifties, all of the most important ideas surrounding problem solving had been determined, almost every currently important application was in the research agenda, and the basic notions for the creation of programs to simulate search processes (satisficing, heuristics) had been set up. The basic pillar of artificial intelligence was identified in parallel with the first studies on the psychology of problem solving.

5. Increasing difficulties

Simon's publication list includes more than a thousand works. To mention only some of them, the most well-known and 'classical' books include *Administrative Behavior* (1947), *Organizations* (1958), co-authored with James G. March, *The Sciences of the Artificial* (1969), *Human Problem Solving* (1972), co-authored with Allen Newell, *Models of Bounded Rationality* (1982), and *Models of Discovery* (1977). It is easy to recognize that the revolutionary ideas he presented during the fifties bore intellectual fruits for the rest of his life. To say that all the scientific production after the fifties incorporates elements of the key discoveries he made in the fifties would not be an exaggeration. Unfortunately, the climate that, despite academic quarrels with the "economists", allowed for a decade of incredibly rich production of innovative ideas went on to change dramatically.

> "By the early 1960s, the Golden Age of organization theory and the behavioral theory of the firm had ended at Carnegie Institute of Technology. […] GSIA came to be dominated by research on sophisticated mathematical techniques in operations research and economics and by neo-classical economic theory. The economists' aborted revolution of 1951 achieved a large measure of success in the 1960s." (Simon, 1996, p. 164).

Ironically enough, the trigger for the final clash came from the kind of collaboration with economists that Simon had explicitly wished for. Simon took part in a study on decision making under uncertainty, conducted jointly with Charles Holt, Franco Modigliani, and John F. Muth (dubbed the HMMS team, with Simon representing the acronym's final letter). The study's aim was to develop mathematical tools to improve inventory control systems for production planning at a Pittsburgh Plate Glass Company plant. It was in this context of a concrete study of empirical data that Simon developed his early notions of dualism between routine and innovative behavior and the notion of "satisficing" behavior, from where the two opposite ideas of rational expectations (Muth, 1961) and bounded rationality flowed.

> "It is not without irony that bounded rationality and rational expectations, two of the major proposals after Keynes for the revision of economic theory (game theory is a third), though entirely antithetical to each other, were engendered in and flourished in the same small business school at almost the same time.
> Not only did they flourish, but they were represented, along with Keynesian theory, in a four-man team that worked closely and amicably together for several years on a joint research project. The HMMS research team harbored simultaneously two Keynesians (Modigliani and Holt), the prophet of bounded rationality (Simon), and the inventor of rational expectations (Muth), the previous orthodoxy, a heresy, and a new orthodoxy." (Simon, 1996, p. 250)

Simon's friendship with the HMMS team did not prevent the rise of a hostile mentality within the GSIA.

> "Over time, a coalition of neoclassical economists and operations research specialists came to dominate the GSIA senior policy committee, making decisions that produced a growing imbalance in the composition of the faculty. Although I had never thought I lacked sympathy with mathematical approaches to the social sciences, I soon found myself frequently in a minority position when I took stands against what I regarded as excessive formalism and shallow mathematical pyrotechnics. The situation became worse as a strict neoclassical orthodoxy began to gain ascendancy among the economists. It began, oddly enough, with Jack Muth. […]
> The rational expectationists, and the neoclassical mathematical economists generally, gradually made GSIA less and less congenial to me. To oppose the trend and secure more tolerance for other points of view, I would have had to devote most of my time to the politics of GSIA, which was not where my interests then lay. It is not clear whether I would have won the struggle had I undertaken it." (Simon, 1996, pp. 249-250).

Around 1970, Simon moved his office to the psychology department. This accelerated his commitment to artificial intelligence and cognitive psychology. His research thus shifted to understanding the various mental abilities essential for human action – memorization, categorization, judgment, and induction – with increasingly intense experimental scrutiny. The buildup of formal models of chess playing, still considered one of the most important roots of artificial intelligence, was accompanied by experiments on the psychology of chess players, experiments that Simon developed for many decades, with many coauthors.[7] He developed the idea of the "human solver" and sought to clarify the shortcomings of the traditional theory of decision making in light of psychological features of human reasoning.

6. The clash between rational expectations and bounded rationality

Two opposite intellectual histories originated from the HMMS team's ideas, spreading their effect for more than half a century, each of them conquering a dominant intellectual position in different scientific fields and in different decades. On the one hand, the powerful ideas stemming from rational expectations' assumptions, after many decades of dominance, have been criticized from the point of view of both experimental validation and internal consistency. On the other hand, the bounded rationality approach, after migrating from economics to psychology and artificial intelligence, moved back to economics through the rapid development of the behavioral approach, and is now undergoing a period of strong intellectual expansion. By briefly following the two paths, I will evaluate current state-of-the-art perspectives and what could be presumed for the future.

---

[7] The most significant collaborations are with Fernand Gobet and with William Chase.

6.1. From expected utility theory to a consistent theory of behavioral inconsistencies

The most important critique of rational expectations comes from behavioral economics, a critique that originates from the studies of the experimental anomalies of the expected utility theory. A brief sketch of the most important steps of this critique begins with, as recalled above, Maurice Allais' experiments in 1952 on individual preferences, which showed systematic deviations from theoretical predictions. Allais' experiment shows the failure of one of the theory's axioms: cancellation.[8] Given the climate of religious respect for the theory during the fifties, it took more than twenty years to convince economists that the discrepancy Allais discovered should be seriously considered.

Especially from the mid 1970s onwards, many proposals were put forward to respond to Allais' paradox, all of them attempting to relax or slightly modify the theory's original axioms.[9] None of them found statistical confirmation over the full domain of applicability (Tversky and Kahneman, 1986, p. 88). Therefore, this strategy of taking the anomalies Allais discovered into account did not prove successful.

Moreover, in 1981 and 1986, Kahneman and Tversky showed the failure of another important axiom: invariance. This was the most dangerous type of failure because it meant that individuals react differently based on how the choice is represented.[10]

A growing body of evidence progressively revealed that individuals do not necessarily conform to the theory of decision-making's predictions, but instead seem to depart from them systematically. The

---

[8] In its standard form, expected utility theory is based on four substantive assumptions: cancellation, transitivity, dominance, and invariance (besides the more technical assumptions of comparability and continuity).

[9] Among others, the weighted utility theory, the regret theory, and the disappointment theory.

[10] The most impressive case is the framing effect: a large majority of individuals behave as risk takers when faced with a problem presented in terms of loss, but they become risk averse when the same problem is presented in terms of gain.

discovery of further categories of violations gave rise to a curious dilemma; either accept a perfect and elegant theory yielding largely false predictions of human behavior, or move toward a new theory that could essentially account for the behavioral inconsistencies. Kahneman and Tversky's approach goes in this direction—instead of limiting their work to modifying certain axioms, they reframe the problem, referring to the mental processes involved, i.e. the process of human *reasoning*. In short, the objective was to create a consistent theory of behavioral inconsistencies. A complete theory of human reasoning does not exist thus far, but over many decades, numerous scholars set the basic pillars of the analysis of the reasoning process, with close links to Simon and his colleagues' research on the behavior of chess masters.

Simon maintains that, in the process of acquiring their skill, chess players store chunks in their long-term memory that correspond to the patterns of pieces. Their recollections from long-term memory is fast and automatic during the match, forming the basis for the conscious process of symbolic manipulation of recalled mental items. This dualism between the unconscious and deliberate aspects of the thinking process has been further explored over the past years, also outside of the context of chess, by many authors, and with particular intensity by Schneider and Shiffrin (1977). They furthered Simon's discoveries, deepening the analysis of dualism in reasoning and better qualifying the process. Thinking is once again considered as emerging from an interaction between two kinds of processes, one of which is 'automatic' or 'intuitive' and allows the retrieval of items from long-term memory, occurring without the subject's control and without necessarily demanding attention. The second process is 'controlled' or 'deliberate'—a conscious elaboration of thoughts that requires attention and is capacity-limited, serial, and controlled by the subject.

From this distinction emerged two different lines of research. On the one hand was the cognitive study of human problem solving and human discovery. According to Simon, solving a problem requires a selective search that relies on pattern recognition and heuristics, two elements that require automatic recollection from long-term memory prior to conscious deliberation. In this context, the interaction between automatic and deliberate thinking explains, for example, the chess master's superiority over a novice. The master has the advantage of automatically using heuristics accumulated through years of constant practice (Chase and Simon, 1973, p. 56). In the same vein, the more a novice learns and memorizes new strategies, the less mental effort he needs to play.

On the other hand, the interaction between automatic and deliberate thinking is used to explain deviations from rationality. Kahneman and Tversky conducted numerous studies about human heuristics using a great variety of choice problems, exploring the interferences between the items stored in long-term memory and the process of conscious deliberation. Their basic finding is that the automatic process can override deliberation with intuitive answers, undermining the workings of the deliberate process.

Human decision-making, then, may be biased because of the interference of intuition, resulting from the domination of inappropriate heuristics over reasoning, and correcting this type of error requires time and mental effort. According to Kahneman, "the acquisition of skills selectively increases the accessibility of useful responses" (Kahneman, 2003b, p. 700). Chess provides a clear example. The more expert the chess player, the more sophisticated and efficient the heuristic that emerges when faced with the same distribution of pieces on a given chessboard.

All this forms the major premises for a theory of reasoning that could explain the inconsistencies of human behavior—Kahneman and Tversky's prospect theory. Whatever the evaluation of prospect theory's impact, I believe that its general direction is well inscribed into the path created by Simon's ideas.[11] As we will see, it also constitutes the basic pillars for a critique of rational expectations.

---

[11] Tversky and Kahneman clearly recognize the consistency of their approach with Simon's conception of bounded rationality (Tversky and Kahneman, 1986, pp. 272-273).

6.3 *Paths in finance: the heterogeneity of expectations*

One field in which bounded rationality's impact has been remarkable is on the analysis of financial markets, which had been relying on the theory of rational expectations.

Let us briefly recapitulate the problem's main features. According to Friedman, rational individuals are supposed to make decisions consistent with optimization principles such as expected utility or expected profit maximization. Markets converge to equilibrium thanks to the actions of rational traders, who are supposed to correctly evaluate prices. The problem lies in evaluating the 'right' prices. In financial markets, a security's price equals its "fundamental value". This is the discounted sum of expected future cash flows where, in forming expectations, investors correctly process all available information. Then, according to Friedman, as soon as there is a deviation from the fundamental value, an attractive investment opportunity is created and rational traders will discover the opportunity, thereby correcting the mispricing.

In this framework, many different kinds of expectations can be attributed to individuals (adaptive, extrapolative, and so on). The implication is that a variety of different expectations about future events are possible, and if agents display a large variety of different expectations, only a fraction of them will be realized. As a consequence, a large number of traders are supposed to fail in their prediction, and the traders who are successful in correcting mispricing are the ones whose predictions are revealed to be rational *ex post*.

Muth (1961) assumed that, on average, rational agents' expectations must be consistent with the events that follow their choices, i.e. agents do not make systematic forecasting errors. He then added more rationality to Friedman's approach, while Simon did the opposite. Rational expectations provided an elegant way to exclude the previous variety of forecasting rules and make psychology irrelevant for economic modelling.

Since its introduction by Muth and its use in macroeconomics by Lucas, the rational expectations hypothesis has been the dominating expectation formation paradigm in economics for many decades. In finance, Muth's proposal was supplemented by Fama's (1970) efficient markets hypothesis (EMH). If markets were not efficient, then there would be unexploited profit opportunities, which would be exploited by rational arbitrage traders.

> "In an efficient market, there can be *no forecastable structure* in asset returns, since any such structure would be exploited by rational arbitrageurs and therefore disappear" (Hommes, 2006, p. 1113).

This means that if expectations are rational, then arbitrage becomes immediate, since traders not only use all available information but also correctly forecast prices.[12] Under these conditions, the efficient markets hypothesis predicts that market behavior can be described by a random walk process. Rational traders would immediately adjust the price to the fundamental value; therefore there is no more room for predicting price movements generated by mispricing. The only element that would have an impact on prices would be external news and shocks, which are not predictable by definition. Thus, the EMH can be verified by testing the randomness of price movements. Tests of the efficient market model are numerous, and not all results have fully confirmed the assumption of randomness, paving the way for contrasting opinions.

By the beginning of the twenty-first century, the intellectual dominance of the efficient market hypothesis had become far less universal. The skeptical views were either from analytical critiques or the discovery of anomalies in experimental data. The most important contributions in the latter area are probably due to Shiller, who notes:

> "[t]he anomalies that had been discovered might be considered at worst small departures from the fundamental truth of market efficiency, but if most of the volatility in the stock market were unexplained, it would call into question the basic underpinnings of the entire efficient markets theory. […] The evidence regarding excess volatility seems, to some observers at least, to imply that changes in prices occur for no

---

[12] The idea that speculative asset prices such as stock prices always incorporate the best information about fundamental values and that prices change only because of good, sensible information meshed very well with theoretical trends of the time.

> fundamental reason at all, that they occur because of such things as 'sunspots' or 'animal spirits' or just mass psychology." (Shiller, 2003, p. 84).

The analytical critiques relate to the incompleteness of Muth's assumptions. Put in a very simplified way, the debate clarified that a pure market with rational traders who have rational expectations cannot exist. Should all traders have perfect expectations about a given security (and given conditions of common knowledge), then all of them would share the same opinion, and all of them would be willing to buy or sell the same security, thus closing the market (Aumann, 1976). The missing information or noise makes financial markets possible, but it also makes them imperfect (Black, 1986, p. 530).[13]

The amazing conclusion is that heterogeneity of opinions is essential for trading to take place, which fits well with the context of bounded rationality. The idea that the fundamental value of a security can easily and immediately be computed seems, then, to be unviable. This moves us back to Keynes, who posited that investors mainly make decisions based on short-term evaluations that are largely dependent on psychological elements.

> "Investment based on genuine long-term expectation is so difficult as to be scarcely practicable. He who attempts it must surely lead much more laborious days and run greater risks than he who tries to guess better than the crowd how the crowd will behave; and, given equal intelligence, he may make more disastrous mistakes" (Keynes, 1936, p. 157).

According to Keynes, it is hard to compute an objective measure of market fundamentals and costly to gather all relevant information, if the latter is even possible. This has been largely developed in the field of *behavioral finance* and subjected to many experimental validations (Barberis and Thaler, 2003). Moreover, new models based on the heterogeneity of trader expectations have been introduced to

---

[13] In a similar vein, Grossman and Stiglitz (1976) claim that because information is costly, prices cannot perfectly reflect the available information, since if they did, those who spent resources to obtain it would receive no compensation. Then, prices must only partially reflect the information of arbitrageurs, so that those who expend resources to obtain information do receive compensation.

take into account the distinction that exists in the real business community between ‘fundamentalist’ and ‘chartist’ traders. Cars H. Hommes framed the situation well:

> “[r]ecently a number of structural asset pricing models have been introduced, emphasizing the role of heterogeneous beliefs in financial markets, with different groups of traders having different expectations about future prices. In all these heterogeneous agent models different groups of traders, having different beliefs or expectations, coexist. Two typical trader types can be distinguished. The first are rational, ‘smart money’ traders or fundamentalists, believing that the price of an asset is determined completely by economic fundamentals. The second typical trader type are ‘noise traders’, sometimes called chartists or technical analysts, believing that asset prices are not completely determined by fundamentals, but that they can be predicted by simple technical trading rules based upon patterns in past prices, such as trends or cycles.” (Brock and Hommes, 1998, p. 1235).

7. Paths in consumer choice and political choice [14]

Beyond financial choices, one of the potentially most fruitful developments of bounded rationality in the years to come is its application to consumer choice and political choice. Both types of choice are characterized by behavior in conditions of low competence. Consumer behavior is, in fact, biased in many different ways, and an increasing body of research shows that, thanks to these biases, consumers’ preferences may be subtly manipulated through advertising.

Schumpeter neatly depicted this process in *Capitalism, Socialism and Democracy*, where the limits of human rationality are highlighted in a way that is impressively akin to Simon’s view. Schumpeter considers the limits of rationality as a universal trait of human thinking, which characterizes both consumer choice on the market and voting decisions in the political arena.

---

[14] This paragraph summarizes the ideas expressed in Egidi (2017).

> "Economists, learning to observe their facts more closely, have begun to discover that, even in the most ordinary currents of daily life, their consumers do not quite live up to the idea that the economic textbook used to convey. On the one hand, their wants are nothing like as definite and their actions upon those wants nothing like as rational and prompt. On the other hand they are so amenable to the influence of advertising and other methods of persuasion that producers often seem to dictate to them instead of being directed by them." (Schumpeter, [1943] 2003, p. 257).

Schumpeter vividly describes a process that has been empirically studied in depth in the last three decades in cognitive psychology. In the domains in which the competence of an individual is very limited, the semantic content of the thought may override the individual's rationality. Prior knowledge and, in particular, social rules and norms, may dumb down the ability to make inferences. This phenomenon is demonstrated by two key experiments in cognitive psychology—the belief-bias effect (Evans, 2003) and the Wason selection task (Wason, 1968).

Therefore, prior beliefs, prejudices, and social norms may have a major role in either hindering or easing the reasoning process. This fragility of the human capacity for reasoning can only be corrected very slowly because the recollection of items that could hinder or foster the responses to given problems cannot be consciously controlled, i.e. memorized items have different degrees of accessibility. Consequently, the only way to modify accessibility is through a slow process of critical reasoning and learning through which the influence of prior beliefs and prejudices can be removed.

According to Schumpeter, individuals show limited rationality in some areas that are removed from their professional competence, and this is particularly common in the area of political choice.

> "The old theory of democracy as formulated in the seventeenth and eighteenth centuries presupposes degrees of awareness of one's interests, clearness of ends, rationality in the perception and use of means and, most important of all, accessibility to rational argument which are altogether unrealistic. A reformed theory of democracy could still use, to a considerable extent, rational schemata, but it would have

to drop, not wholly but also to a considerable extent, the hypothesis of conscious rationality." (Schumpeter, 1984, p. 585).

The fact that individuals use a low level of competence in the fields that are not directly connected with their everyday activity has been the basic assumption of the modern theory of democracy, developed by Schumpeter in 1950. By contrasting the pretense of Olympian rationality, on which the traditional French theory of democracy is founded, he claims that exactly because of the modest competence and effort, on average, in citizens' rationality, democratic institutions must exist. In his view, democracy can, in principle, guarantee that citizens really choose in accordance with their goals and beliefs, despite their modest competence, through the existence of parties and leaders. Political parties are charged with the responsibility of serious informational and persuasive work. Moreover, the advertising process may have strong distortive effects in the case of consumer behavior, however on the other hand, in a well-functioning democracy normal and institutional conflict among the different parties (and connected programs and ideologies) may allow citizens to select the political proposals that are more fitting to their needs and values.

Schumpeter's broad view of rationality contains in nuce and in a simplified way certain principles of rationality that are strikingly similar to the ones that preside over Simon's theory of bounded rationality. With his conceptual distinction between high and low degrees of conscious rationality, he can encompass two different kinds of behavior—the innovative and the routinized. The first form of behavior arises from situations of unpredictability, uncertainty, and complexity. Accordingly, Schumpeter treats it as a problem of research and innovation. His description of innovative activity fits with Simon's notion of "procedural rationality" that describes the research and decision-making process of an expert in conditions of uncertainty.

7. Concluding remarks

I believe that Simon's vast intellectual contribution to science can lead to a unified view within the social sciences, based on a new understanding of human cognitive activities. The core of his views has produced radical mutations in ways by which we interpret human behavior, and caused the incredible expansion of the fields of behavioral sciences, artificial intelligence and the sciences of organization; it promises great potential for further intellectual achievements.

References

Allais M. (1952), "Le comportement de l'homme rationnel devant le risque: Critique des postulats de l'école Américaine", *Econometrica*, vol. 21 n. 4, pp. 503-546.
Amarel S. (1968), "On the Representations of Problems of Reasoning about Actions", in Michie D. (ed.), *Machine Intelligence 3*, New York: American Elsevier, pp. 131-171.
Aumann R.J. (1976), "Agreeing to Disagree", *The Annals of Statistics*, vol. 4 n. 6, pp. 1236-1239.
Barberis N., Thaler R. (2003), "A Survey of Behavioral Finance", in Stoll H.R, Constantinides G.M., Harris M., Stulz R.M. (eds.), *Handbook of the Economics of Finance*, volume 1, Amsterdam: North Holland, pp. 1053-1128.
Barnard C. (1938), *Functions of the Executive*, Cambridge: Harvard University Press.
Becker M.C., Lazaric N., Nelson R.R. (2005), "Applying Organizational Routines in Understanding Organizational Change", *Industrial and Corporate Change*, vol. 14 n. 5, pp. 775-791.
Black F. (1986), "Noise", *The Journal of Finance*, vol. 41 n. 3, pp. 528-543.
Brock W.A., Hommes C.H. (1998), "Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model", *Journal of Economic Dynamics and Control*, vol. 22 n. 8, pp. 1235-1274.
Chase W.G., Simon H.A. (1973), "Perception in Chess", *Cognitive Psychology*, vol. 4 n. 1, pp. 55-61.
Cohen M.D., Bacdayan P. (1994), "Organizational Routines Are Stored as Procedural Memory: Evidence from a Laboratory Study", *Organization Science*, vol. 5 n. 4, pp. 554-568.
Cohen M.D., Burkhart R., Dosi G., Egidi M., Marengo L., Warglien M., Winter S.G. (1996), "Routines and Other Recurring Action Patterns of Organizations: Contemporary Research Issues", *Industrial and Corporate Change*, vol. 5 n. 3, pp. 653-698.
Cyert R.M., Simon H.A., Trow D.B. (1956), "Observation of a Business Decision", *Journal of Business*, vol. 29 n. 4, pp. 237-248.

Dosi G., Nelson R.R., Winter S. (2001), *The Nature and Dynamics of Organizational Capabilities*, Oxford: Oxford University Press.
Egidi M., Narduzzo A. (1997), "The Emergence of Path Dependent Behaviors in Cooperative Contexts", *International Journal of Industrial Organization*, vol. 15 n. 6, pp. 677-709.
Egidi M. (2017), Schumpeter's picture of economic and political institutions in the light of a cognitive approach to human behavior *Journal of Evolutionary Economics vol.* 27,pp.139–159
Evans J. (2003), "In Two Minds: Dual Process Accounts of Reasoning", *Trends in Cognitive Sciences*, vol. 7 n. 10, pp. 454-459.
Fama E.F. (1970) "Efficient Capital Markets: A Review of Theory and Empirical Work", *The Journal of Finance*, Vol. 25, No. 2, pp. 383-417
Feldman M.S., Pentland B.T. (2003), "Reconceptualizing Organizational Routines as a Source of Flexibility and Change", *Administrative Science Quarterly*, vol. 48 n. 1, pp. 94-118.
Frenken K., Marengo L., Valente M. (1999), "Interdepencies, Nearly-decomposability and Adaptation", in Brenner T. (ed.), *Computational Techniques for Modelling Learning in Economics*, Boston: Kluwer Academics, pp. 145-165.
Friedman M. (1953), "The Methodology of Positive Economics", in (id.), *Essays in Positive Economics*, Chicago: University of Chicago Press, pp. 3-34.
Grossman S.J., Stiglitz J.E. (1976), "Information and Competitive Price Systems", *The American Economic Review: Papers and Proceedings*, vol. 66 n. 2, pp. 246-253.
Hommes C.H. (2006), "Heterogeneous Agent Models in Economics and Finance", in Tesfatsion L., Judd K.L. (eds.), *Handbook of Computational Economics,* volume 2, Amsterdam: North Holland, pp. 1109–1186.
Kahneman D. (2003a), "Maps of Bounded Rationality: A Perspective on Intuitive Judgment and Choice", *The American Economic Review*, vol. 93 n. 5, pp. 1449-1475.
Kahneman D. (2003b), "A Perspective on Judgment and Choice – Mapping Bounded Rationality", *American Psychologist*, vol. 58 n. 9, pp. 697-720.
Keynes J.M. (1936), *The General Theory of Unemployment, Interest and Money*, London: Macmillan.
Knoblock C.A. (1990), "A Theory of Abstraction for Hierarchical Planning", in Benjamin D.P. (ed.), *Change of Representation and Inductive Bias*, Boston: Kluwer, pp. 81-104.
Knoblock C.A. (1991), "Search Reduction in Hierarchical Problem Solving", *Proceedings of the Ninth National Conference on Artificial Intelligence*, pp. 686-691.
Korf R.E. (1980), "Toward a Model of Representation Changes", *Artificial Intelligence*, vol. 14 n. 1, pp. 41-78.
Levinthal D.A., March J.G. (1993), "The Myopia of Learning", *Strategic Management Journal*, vol. 14 n. S2, pp. 95-112.
Levitt B., March J.G. (1988), "Organizational Learning", *Annual Review of Sociology*, vol. 14, pp. 319-340.
March J.C., Simon H.A. (1958), *Organizations*, New York: John Wiley.

Marengo L. (2015), "Representation, Search, and the Evolution of Routines in Problem Solving", *Industrial and Corporate Change*, vol. 24 n. 5, pp. 1-30.
Marengo L., Pasquali C., Valente M. (2005), "Decomposability and Modularity of Economic Interactions", in Callebaut W., Rasskin-Gutman D. (eds.), *Modularity: Understanding the Development and Evolution of Complex Natural Systems*, The Vienna Series in Theoretical Biology, Cambridge (MA): MIT Press, pp. 835-897.
Muth J.F., (1961), "Rational Expectations and the Theory of Price Movements", *Econometrica*, vol. 29 n. 3, pp. 315-335.
Nelson R.R., Winter S.G. (1982), *An Evolutionary Theory of Economic Change*, Cambridge (MA): Harvard University Press.
Newell A., Simon H.A. (1972), *Human Problem Solving*, Englewood: Prentice Hall.
Polya G. (1945), *How To Solve It*, Princeton: Princeton University Press.
Schneider W., Shiffrin R.M. (1977), "Controlled and Automatic Human Information Processing: I. Detection, Search, and Attention", *Psychological Review*, vol. 84 n. 1, pp. 1-66.
Schumpeter J.A. (1943), *Capitalism Socialism and Democracy*, New York: Harper & Brothers; 2nd edition 2003, London: Routledge.
Schumpeter J.A. (1984), "The Meaning of Rationality in the Social Sciences", *Zeitschrift für die gesamte Staatswissenschaft – Journal of Institutional and Theoretical Economics*, vol. 140 n. H4, pp. 577-593.
Shiller R.J. (2003), "From Efficient Markets Theory to Behavioral Finance", *Journal of Economic Perspectives*, vol. 17 n. 1, pp. 83-104.
Simon H.A. (1947), *Administrative Behavior*, New York: Macmillan.
Simon H.A. (1963), "Problems of Methodology – Discussion", *The American Economic Review: Papers and Proceedings*, vol. 53 n. 2, pp. 229-231.
Simon H.A. (1969) *The Sciences of the Artificial,* Cambridge (MA): The MIT Press
Simon H.A. (1982), *Models of Bounded Rationality,* Cambridge (MA): The MIT Press.
Simon H.A. (1977), *Models of Discovery* , Reidel, Dordrecht, Holland,
Simon H.A. (1996), *Models of My Life*, Cambridge (MA): The MIT Press.
Simon H.A. (2002), "Near Decomposability and the Speed of Evolution", *Industrial and Corporate Change*, vol. 11 n. 3, pp. 587-599.
Tversky A., Kahneman D. (1986), "Rational Choice and the Framing of Decisions", *The Journal of Business*, vol. 59 n. 4, pp. S251-S278.
Wason P.C. (1968), "Reasoning about a Rule", *Quarterly Journal of Experimental Psychology*, vol. 20 n. 3, pp. 273-281.
Zollo M., Winter S.G. (2002), "Deliberate Learning and the Evolution of Dynamic Capabilities", *Organization Science*, vol. 13 n. 3, pp. 339-351.