

# Learning, Rationality and Identity Building - From Simon To Varela<sup>1</sup>

**David CAYLA**

Contact : [dcayla@free.fr](mailto:dcayla@free.fr)

**IMRI**

Université Paris-Dauphine  
Place du maréchal de Lattre de Tassigny  
75775 Paris CEDEX 16

**CEA**

DRT/DPSE

Bat. 446

91191 GIF SUR YVETTE CEDEX

*September 2008*

## **Abstract**

This paper focuses on the link between the economic conceptions of rationality and learning. Traditionally, most economists believe that learning is just a way for agents to become fully rational. But being fully rational cannot describe a process, for there is only one way to be rational in the economic sense of the term. Therefore, what economists have in mind is not the *process* of learning, but the *result* of learning: ‘a fully rational agent’. Heterodox rationality conceptions such as the Simonian model of bounded rationality seem more compatible with the idea of learning. Bounded rationality implies that agents may act differently to the same stimulus; it is therefore compatible with the idea of diversity, one of the foundations of the evolutionary logic. But following Simon, learning should not be considered as a creative process that allows a lot of diverse answers. If diversity exists in the agents’ behaviors, the way they learn appears to be unique. As a consequence, learning should decrease the strength of the selection forces, both processes being contradictory (Dosi *et al.* 2003). Our paper aims to overcome this contradiction by showing how intentionality and identity, and more broadly Francisco Varela’s ‘enaction’ theory, can help to invent a concept of ‘rational learning’ that is compatible with the evolutionary logic.

---

<sup>1</sup> *This paper was presented at the IAREP/SABE meeting at LUISS in Roma, September 3-6 2008. The author is very grateful to Eshien Chong and Renelle Guichard for their helpful comments.*

## 1. Introduction

In the introduction of a long review devoted to learning in economics, Dosi, Marengo and Fagiolo wrote that ‘learning [...] is an ubiquitous characteristic of most economic and, generally, social environments, with the remarkable exception of those postulated by the most extreme forms of economic modelling, such as those assuming rational expectations or canonical game-theoretic equilibria’ (Dosi *et al.* 2003: 3). Truly, if the notion of learning is essential to every study on human environments, for a long time, economists have been using it with infinite caution. The study of learning did not appear as a priority and most of them would have been delighted to leave this question to psychology or other social sciences. To take the words of Lazaric *et al.* ‘Learning in itself does not interest economists, but it is its consequences or its economic expressions that do’ (Lazaric *et al.* 1995: ix).

Yet, as economic sciences get closer to other social sciences, especially business, or when they try to go deeper into the psychological or behavioral foundations of the agents they study, it becomes more and more difficult to dismiss this question. In this regard, heterodox approaches based on Nelson and Winter (1982) seminal work were fast to acknowledge the necessity to take learning into account in evolutionary environments.

This was not without consequences. Studying learning is indeed impossible without interrogating our rationality conceptions (Kirman and Salmon 1995). Aware of this difficulty, evolutionary economists have chosen to found their theory on the Simonian bounded rationality model instead of the perfect rationality model used in standard economic conceptions. But this strategy raised new difficulties. In the Simonian approach, learning is a pure adaptive process and comes down to the memorization of information given by the environment. Two agents faced with the same environment will therefore learn the same thing and their behaviors will progressively converge.

Unfortunately, this idea is contradictory with an evolutionary fundamental principle: the diversity of responses. As Stanley Metcalfe explicitly emphasizes, ‘what would kill the evolutionary argument stone dead would be if all units of selection adapted their behaviour in identical fashion to the appropriate signals. Then we would have uniform responses, no variety and no evolution’ (Metcalfe 2005: 414). In the same logic, Dosi *et al.* (2003: 62) argue that selection and learning are contradictory, for in a ‘pure learning’ process, ‘all agents make the best use of the available information, are endowed with identical information-processing algorithms, etc’. As a result, there is no diversity in the way they behave, and selection can play no role in the evolution process.

The goal of this paper is to overcome the contradiction between selection and learning by proposing a ‘post-Simonian’ approach of rationality that is compatible with the idea of a subjective learning. Indeed, if we can show that there are several ways to learn, several paths to follow, then selection can occur at any step of the learning process. If the agents’ behaviors do not converge while they learn, then there is no reason to think that learning weakens the selection forces in a population.

In order to emphasize the diversity of the learning processes, we will show that the logic of learning is based on a *conscious choice* that may look like a ‘bet’. Relying on Fransisco Varela’s theory of cognitive sciences, a ‘creative learning’ approach will be developed. In this approach, learning is considered not to be equivalent to acquiring true information about the environment, but as the deepening of an individual *coherence* that stands at the interface between the world reality and the *identity* of learning organisms.

Our reasoning will follow three steps. The next section will briefly show that the notion of learning is not compatible with the idea of perfect rationality. Section 3 will be devoted to the study of the Simonian bounded rationality model and to its limits. At last, the fourth section will develop a new approach of the relation between behavior and identity that will allow us to propose this new learning model.

## **2. Learning and rational behavior**

If economists are interested in rationality models (and especially in perfect rationality models), it is because they are useful to establish what kind of decision an agent will take and what its behavior will be. In this regard, Walliser (2000: 72) describes as follow the three characteristics of rational decision models:

- No social dynamic: The rational model implies that only individuals decide, without the influence of a group. As a consequence, the social network can be split up in a vast number of individual decision centres.
- Nor coherence neither complementarities in behaviors. This assumption implies that all actions of an agent can be divided into individual actions and can be analyzed separately.
- The decision process is made up of three successive steps: i) information gathering; ii) deliberation; iii) implementation.

Yet, this model raises some specific difficulties for the study of learning. Contradicting the first characteristic Walliser raised, most learning theories are based on the capacity individuals have to imitate or to be influenced by others' experience (*cf.* Simon's (2005) 'docility' concept or the 'social learning' approach of Ellison and Fudenberg 1993). The second difficulty comes from the second assumption, which postulates that the successive actions of an individual may be split up in independent unite, while learning implies a cumulative dimension and dependence between the different steps of a learning process. At last, the third difficulty deals with the hypothesis of a strict linearity between preferences and behavior. It prevents one from considering the eventually endogenous character of preferences.<sup>2</sup>

Besides the criticisms that can be levelled against the standard decision model, one can wonder if the 'perfect rationality' concept is adapted to any conception of learning. As a matter of fact, if agents are seen as being perfectly rational, if they have a perfect knowledge of their environment, and if they enjoy an infinite computing capacity, we hardly see how (and what) they could still learn (Kirman and Salmon 1995).

Walliser (1995) distinguishes two types of rationality: on one hand the *cognitive rationality* assumes the appropriateness between the agent's beliefs and the information he received. It is used to build the agent's expectations. On the other hand the *instrumental rationality* assumes the appropriateness between the agent's opportunities and his preferences. It is used to help the agent develop a strategy that matches its expectations.

In the case where an individual has a perfect cognitive rationality and a perfect instrumental rationality, it is considered to be 'substantially' rational. In Simon's sense, that means 'that behavior is objectively rational in relation to its total environment, including both present and future environment as the actor moves through time' (Simon 1986: 210).

This kind of rationality implies a very specific form of learning that Walliser (1998) calls 'eductive process' (the expression is taken from Binmore 1987). In an eductive process, agents are perfectly informed about their environment and they can treat perfectly this information. As a consequence, each agent shares the same knowledge and the same representation. This kind of process logically implies a perfect prediction about the world and about other agents' behaviors. Everyone will therefore instantaneously be in equilibrium, without any dialogue.<sup>3</sup>

---

<sup>2</sup> Dosi *et al.* (2003) estimate that it is a crucial assumption of the standard economic model: 'Endogenous preference may be often driven by the attempts to reduce regret and cognitive dissonance [...] ...if you cannot be with the one you love, love the one you are with... !' (Dosi *et al.* 2003: 15).

<sup>3</sup> Walliser (1998: 72) speaks of an equilibrium guided by a 'Lewisian hand', that is by the existence of a 'common knowledge'.

Is this rationality model compatible with the notion of learning? It depends on the conception we chose to adopt about learning. For Pierre Garrouste, the educative process means that learning is assimilated to a mere information gathering. As he emphasizes, in this kind of process ‘individuals are not really learning, they are only modifying their behavior due to the information they gather’ (Garrouste 2003: 305). This kind of learning is similar to Bateson’s (1972) learning zero: An automatic learning that does not imply any cognitive change in the representation systems.

In other words, it seems difficult to speak of a *process*, for the changes appear instantaneously, in the absence of dialogue as well as introspection. Besides, and it is quite a paradox, this learning model tends to reduce to its simplest terms the agent’s cognitive dimension, suppressing the deliberation step proposed by Walliser. Garrouste (1999: 141) evokes the idea of a ‘maximizer automaton’, whereas Dibbiagio (1999: 121) speaks about a ‘Pavlovian decision process of stimuli-response style’.

#### *Learning: a march towards perfect rationality?*

If the concept of perfect rationality is not adapted to describe the behavior during learning, it may be integrated to learning models as a long term perspective. The idea proposed by several economists is that perfect rationality models may help to understand, not the real behavior of agents, but where learning tends to push them. Learning in this sense would be analyzed as a ‘march toward perfect rationality’.<sup>4</sup> In this context, learning can be a way to found the perfect rationality assumption. But this foundation is still problematic. If learning is a march toward rationality, it means that as soon as perfect rationality is obtained, individuals do not learn anymore. In other words, learning and perfect rationality are still in contradiction. Either individuals learn, or they are perfectly rational.

In a more fundamental way, the idea that individuals may become progressively rational means that it is possible to understand the rationality acquisition as a cumulative process. Yet, there is only one way to be rational by definition. Thinking that it is possible to rank individuals in respect to their learning level (*i.e.* to their rationality degree) is very troublesome. On which criteria shall we determine that such agent is behaving in a ‘more’ or ‘less’ rational way than

---

<sup>4</sup> For Börgers learning theories ‘...describe how economic agents adjust their behaviour over time, and how, after agents have gained experience, their behaviour may become rational in the economists’ sense of the word’ (Börgers 1996: 1).

another? Because rationality is a binary concept, it seems very difficult to use it in order to describe an evolving process.

Finally, this approach of learning is very linear. Learning, in this conception, is seen as an objective process that would conduct the representations of an individual closer to 'reality'. In other words, only one kind of learning model would exist, and there would be only one optimal behavior to converge. The idea of *mistake* in learning is taken off the analysis.

### 3. Learning and bounded rationality

In economics, evolutionary models were the first to fully integrate the Simonian notion of bounded rationality in the core of their theory. Simon believed that the rationality models must be founded on an empirical basis and must be compatible with the ones that are used in other social sciences. As he says, 'in economics, rationality is viewed in terms of the choices it produces; in the other social sciences, it is viewed in terms of the processes it employs' (Simon 1986: 210). Simon proposes therefore a process-based rationality, which he calls *procedural rationality*, as opposed to economists' *substantive rationality*. This approach is set up on the idea that individuals do not base their choices from the world as it objectively is, but from their subjective perception of it.

If [...] we accept the proposition that both the knowledge and the computational power of the decision maker are severely limited, then we must distinguish between the real world and the actor's perception of it and reasoning about it. That is to say, we must construct a theory (and test it empirically) of the processes of decision. Our theory must include not only the reasoning processes but also the processes that generate the actor's subjective representation of the decision problem, his or her frame (Simon 1986: 211).

This observation does not mean that individuals act irrationally. Rubinstein (1998: 21) emphasises the 'strategic' aspect of procedural rationality and distinguishes it from an 'impulsive' action. Similarly, for Simon, 'everyone agrees that people have reasons for what they do' (Simon 1986: 209). Talking about 'bounded rationality' does not imply irrationality, but only means that rational actions are taken in the specific context of particular representations.

Going deeper into Simon's reasoning with Walliser's typology, we find that the agent's rationality is not only bounded in the way he makes decisions (bounded in his instrumental rationality), but also in the knowledge and the perception it has of the world (bounded in his cognitive rationality).

With this distinction, two kinds of learning are possible. A first kind may be developed in order to improve the cognitive content of representations. By acquiring information and knowledge, an individual may improve his representation models and may take his knowledge closer to the real state of his environment. A second kind of learning can be developed in order to improve the agent's instrumental rationality. Here, the individual does not try to change the way he understands his environment, but aims to improve the procedures that guide his choices.

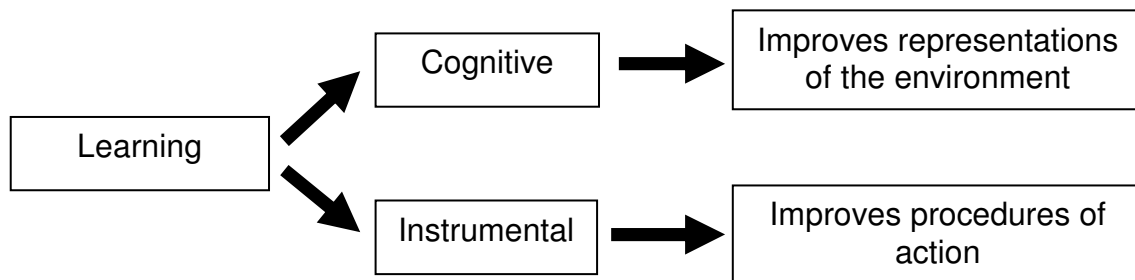


Fig. 1: Two models of learning in bounded rationality

These two models of learning are, of course, complementary. The procedures of action are, as a matter of fact, largely depending on the level of knowledge. Similarly, a change in the decision procedures of an individual may lead him to behave differently, and may help him to develop a cognitive learning from the new experiences he is confronted with.

#### *Limits of the bounded rationality model in the study of learning*

Although the Simonian model of rationality seems to fit better to the study of learning phenomena than the neoclassical one, it has a number of limits.

The first one is related to the way procedures are revised. For Simon, as for Nelson and Winter, it is the 'satisficing' criterion that is put forward. Instead of trying to maximize its utility, the agent tries to reach an objective that he believes is satisfying, and will change his behavior only if this objective is not attained. The question that is left is to understand what kind of learning an agent must develop in case of an unsatisfying result. Does he have to change the deepest procedures (or routines) that determine his behavior, or is it possible to carry out only marginal changes? How will he understand the causes of his failure?

Simon is not very clear in the answers he offers to these questions. The example he gives on the coordination of the American help during the Marshall Plan (Simon 1996: 141-3) is very

revealing about Simon's argument and its limits. Let's briefly recall the problem. When the American government decided to engage a first \$5.3 billion help in order to restore the postwar economy of European nations in 1948, a special agency was created to coordinate this help, the *Economic Cooperation Administration* (ECA). If we are to believe Simon, the agency actions could have been implemented through six different conceptions, and could answer to six different objectives, most of them being contradictory with the others.<sup>5</sup> At its beginning, each conception was present in the agency, but it would have been inconceivable to follow all of them without ruining the coherence of the American policy towards Europe. Finally, Simon says that after a year of functioning, a common representation emerged in the agency 'as it evolved'.

But this example appears more to contradict than to support Simon's view. Simon defends the idea that an adaptive process, by feedback, is superior to a planned process that would be oriented towards final objectives. He gives two reasons for that: First, an adaptive process helps us to keep a maximum of options open for our future choices (*Ibid*: 167); second, only this kind of evolution is compatible with our limited capacity to foretell the consequences of our decisions (*Ibid*: 163). Yet, in the Marshall Plan example, it is precisely the lack of a clear vision that 'almost' (Simon says) 'create thorough confusion in the agency and among its clients' (*Ibid*: 143).

In a nutshell, if you follow Simon's argument, it would have been necessary to organize the American help in the most flexible way, in order not to be trapped in a unique representation that would prevent the progressive adaptation of the Marshall Plan to the 'reality' of the field; but if you follow his example, it was on the contrary necessary to coordinate from the beginning the American help in order to reach one of the six possible objectives, escaping the confusion that may imperil the project.

Let's go on with three remarks. The first one is that the logic that says that adaptive mechanisms based on feedback are more efficient than planned processes is a logic that looks very much like the optimization logic, which Simon rejects. Indeed, if the goal is not to find the 'best' possible organization, but only to select one that is satisfying, then any one of the six representations would be sufficient to implement the Marshall plan.<sup>6</sup> The second remark is

---

<sup>5</sup> The six different conceptions of ECA actions were the following: the commodity screening approach, the balance of trade approach, the European cooperation approach, the bilateral pledge approach, the investment bank approach, and the policy and administration approach.

<sup>6</sup> Simon estimates that every conception had 'rational basis'. He even says explicitly that 'what was needed was not so much a 'correct' conceptualization as one that could be understood by all the participants and that would facilitate action rather than paralyze it' (Simon 1996: 143).

related to the Simonian implicit hypothesis, which assumes that the conception to follow will spontaneously emerge from the confrontation with the 'real things'. This hypothesis seems very much debatable and has never been proved. At last, the third remark is related to the ambiguous character of environmental feedbacks. An adaptive mechanism is efficient (or even possible), only if clear feedbacks, without any ambiguity, are received by the adaptive system. As soon as the environmental feedbacks are open to interpretations or misunderstanding, the system can no longer objectively adapt itself to its environment. This difficulty is even more important if we consider that the feedback interpretation depends on the representations that anyone adopted. In this case, it is impossible that a structure, which brings together different individuals, acting in relation to incompatible representations and objectives, may be able to spontaneously define a common adaptive process. Even if the environment and the problems are the same, they will not be perceived as such. The capacity for agents to agree on a common representation is therefore weakened, instead of strengthened, by the confrontation with reality, as long as everyone find, in the feedbacks, a validation of their own interpretative frameworks.

To sum up, in order to understand how a learning process is being carried out, one needs to understand how an organization or an individual is brought to choose one representation or objective instead of another. In order to escape the ambiguity of a non-choice, it may be better to start from a subjective *a priori* instead of trying to adapt to 'reality'. Learning will therefore follow a 'bet' logic, in which the path to go is chosen *ex ante*.

Adopting this approach in the study of learning implies that the need for the *coherence* of the process is more important than the need for the *optimization* of a behavior to a given environment. This perspective radically change, as we're going to see now, the view Simon adopted in his study of behavior.

### *Are learning processes intentional?*

The Simonian analysis of behavior and learning is based on the refusal of a teleological explanation. Simon disregards the subjective and intentional parts of behaviors. This makes him analyze human behavior like the one of a computer or of an ant. The three following quotations are very revealing as such:

'As we succeed in broadening and deepening our knowledge [...] about computers, we discover that in large part their behavior is governed by simple general laws, that what appeared as complexity in the computer program was to a considerable extent complexity of the environment to which the program was seeking to adapt its behavior' (Simon 1996: 21).

‘An ant, viewed as a behaving system, is quite simple. The apparent complexity of its behavior over time is largely a reflection of the complexity of the environment in which it finds itself’ (*Ibid*: 52).

‘Human beings, viewed as behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves’ (*Ibid*: 53).

This explicit parallel between human behavior and the one of an ant or of a computer leads us to two consequences.

The first one is that the Simonian model assumes that reality is objectively perceived, and therefore that the agents’ behaviors are the direct consequence of the real state of the environment they face. In other words, even if there are limits to their sensorial capacities, there are no systematic cognitive biases in the way individuals see their environment. If human beings behave like an ant, it is because the representation systems upon which they use to act are not the product of a choice or of an intention, but are directly made of what they see. Therefore, the same kinds of behaviors and the same kinds of learning will be adopted by different agents, as long as they are involved in similar situations. The more they face the ‘field’, the more individuals learn and improve their representation systems. These representations will then have the tendency to converge, getting closer and closer to ‘reality’.

The conclusion is very different if learning is considered to be a process that starts from a bet and which coherence is given, not by the objective state of the environment, but by an individual and subjective choice that is made *ex ante*. In this case, agents learn trying to keep a personal coherence over time. This implies that similar individuals will not learn the same thing, even if they are confronted with the same environment, because they will not interpret the feedbacks in the same way. Learning will then not lead representations to converge, but may, on the contrary, maintain or even deepen individual differences.

The second consequence is related to the way learning procedures are changed. Simon sees them as universal and believes that they are imposed upon the behaving systems. So when it needs to revise its behavior in order to face a new situation, the ant only follows other procedures that are, ultimately, parts of its genetic heritage. An ant neither thinks nor deliberates. It is programmed to follow a certain change instead of another. Unless you change its genetic heritage, you cannot change the way an ant learns.

This statement seems however much less obvious for a human being, for one implements a *conscious* action when one decides to undertake a certain kind of learning instead of another. As such, a human being always keeps some control of what he learns. Unlike the ant, he may

*intentionally* break the learning process in which he is engaged, and decide to change his representative schemes. The way an individual learns does not come down to a system of simple rules, exogenously given, but implies a *choice* that has been *ex ante* defined.

Taking into account these two consequences necessitates enriching the Simonian model by adopting another rationality representation that would be compatible with the ideas of *intention* and *choice* in the learning process. This is what we aim to do in the following section.

#### 4. Intentionality and identity in learning

As it was said before, Simon founds his theory of rational behavior on a system of procedures made from environmental feedbacks, and under the constraint of the cognitive and perception abilities of the system. In regard to cognitive sciences, the Simonian approach is called ‘computationist’ (Paulré 2005) or ‘cognitivist’ (Varela 1997). It is based on the assumption that the human brain, like the cognitive system of an ant, or like a computer program, is an information processing system that works sequentially, and that uses items (or symbols), which represent elements of the environment. This approach implies the following assumptions:

- Learning is assimilated to a memorization process (Simon 1996, chap. 4) that is *universal* since it only accumulates two sorts of items: Either information about the state of the world, which will enrich the representation system, or computing capacities, through the memorizations of complex links between items, which will simplify calculations by developing what Simon calls an *intuition* (Simon 1996: 89-90).
- Learning procedures on a cognitive activity is reduced to simple logical operations. If a system behaves (or learns) differently from another, it is either because the objective characteristics of their environments are different, or because the two systems do not have the same cognitive or sensorial abilities.
- Simon believes that the human memory (and therefore learning) is organized into fundamental elements (*chunks*), each of them representing one dimension of the received information (a stimulus). The knowledge of an individual therefore corresponds to the whole chunks that were previously memorized.<sup>7</sup> This conception in

---

<sup>7</sup> Simon evaluates the memorizing time for a chunk to be about 8 second long, and to 50,000 the number of chunks that are necessary to master a discipline like chess (Simon 1996: 66 and 89-90).

which memory is made of partitioned information data does not fit very well with the idea of tacit knowledge.<sup>8</sup>

Our assumption is that this conception of learning is far too restrictive. It does not take into account the intentional dimension of the *action* of learning, and the *choices* that this action necessitates. The goal of this section is then to enrich the Simonian rationality model and to propose a more accurate analysis of human learning phenomena. In the next subsection, the cognitive approaches of Simon and Varela will be analyzed. Then, a ‘subjective’ view of rationality will be developed before we show, in the last subsection, the role and the importance of coherence and identity in learning.

#### *An alternative approach of cognitive sciences*

Cognitive sciences can be defined as the study of the relation between the brain, the mind and the behavior of thinking beings. This relatively new science (Varela 1997 speaks of 1940 as its birth date) is founded on the one hand on the rejection of the behaviorist argument, and on the other hand on the assumption that it is possible to study cognitive activities as such, leaving aside the biological functioning of the brain.

The Simonian cognitivist paradigm corresponds to the social sciences ‘orthodoxy’, as it emerged in the MIT research centers in the 50s. Since then however, new approaches were developed and other theoretical frameworks were invented, like the ‘connexionist’ approach that came out in the 70s. The main assumption of connectionism is that cognitive systems do not work sequentially, treating discrete symbols that represent concrete elements of reality, but react globally to external stimuli. In a connexionist cognitive system, knowledge is not localized, but depends on the global structure of the system, and the cognitive unit does not process the information data the one after the other but in parallel, each datum contributing to transform the whole architecture. The cognitive system is therefore considered as a self-organized dynamic network, in which each part is in coherence with the others.

The connexionist paradigm allows taking into account some mental phenomena that are neglected in the cognitivist approach. First, it can comprehend the tacit dimension of learning. Memory, in that conception, is not considered as a tank of symbolic items that represent concrete elements of the world, but as a ‘sub-symbolic’ system, the meaning of which is

---

<sup>8</sup> This contradiction between the Simonian cognition (in which memory is assimilated to a stock of information data) and Polanyi’s argument on tacit knowledge (which implies a more intuitive memory) is emphasised by several authors (Bessy, 2003; Foss, 2003; Arena and Lazaric, 2004).

written into the interactions of the elements instead of into the elements themselves (Varela 1997: 79). Some knowledge may therefore be diluted in the global organization of the cognitive system, or may be in an exclusive interaction with others (association principle).

The second interest of the connexionist paradigm is to reveal the *interpretation* phenomena. In the connexionist models, human perception is considered to work in relation with the whole cognitive system, instead of in isolation. Varela (1997: 73-5) shows for example that the vision processing unit is less influenced by the optic nerve stimuli, than by the rest of the brain activity. Consequently, information is transformed (*i.e.* interpreted) while it is perceived, since the optic nerve contributes only in minority to the vision processing system.

Another model that challenges the cognitivist principles is the *enaction* approach, which is proposed by Varela. This approach is compatible with the connexionist paradigm and aims to extend it. Its central assumption is that it is possible to study the cognitive processes without the *representation* notion. Varela considers that cognitive sciences have a very occidental conception of the relations between the world and the mind, as most cognitive scientists adopt the dualist view, in which two separated and distinct entities exist. The mind 'conceives', while the world 'is'. Knowledge, in the cognitivist approach, is only the expression of a 'mirror of nature'. The world is predetermined and learning is restricted to an accumulation of information data that corresponds to its real proprieties.

In Varela's approach, on the contrary, the relations between the world and the mind are circular. The world properties are not predefined but 'enacted', they are the product of *action*. In other words, there is no 'pure' perception, every one of them being directed by an *intention*. The Varela's intentionality conception follows the 'Intentionality' notion of John Searle (1983). It designates the capacity of the mind to get the organism connected with the world. To put it differently, it is because it has an intention that an individual is able to direct its actions or its perception toward an actual or possible object.

To sum up, in the enaction approach, the human brain does not only answer predefined questions, as a computer would, but is able to define its own environment: 'the most important faculty of every living cognition is precisely, in a large extent, to ask the pertinent questions that arise at any time in our life. They are not predefined but *enacted*, they are *being emerged* on a background, and the pertinence criterions are given by our common sense, in a way that is always contextual' (Varela 1997: 91).

The enaction theory deeply changes the way rationality and learning are conceived. The main change deals with the ‘creative’ aspect of human cognition. While learning is limited to a memorization process that aims to adapt the organism to a given reality in Simon's conception, Varela views learning as a *construction* that emerges in the relationship between the individual and his environment. A color or a smell are not the objective products of nature, but emerge in the perception individuals have of them.<sup>9</sup>

The enaction approach does not contradict the connexionist models but opens the cognitive sciences to new questions. The first one deals with the questioning of the notion of information. Traditionally, information is seen as an objective property of the environment that would be assimilated (cognitivist paradigm) or interpreted (connexionist paradigm) by the system during the learning process. On the contrary, for Varela, information is nothing else than a ‘modern phlogiston’ (*Ibid*: 11-2). It exists only when it is seen. As he says: ‘information is not pre-established as an intrinsic order, but [...] corresponds to regularities that emerge from the cognitive activities themselves’ (Varela 1997: 122).

In this view, learning has to be considered as a process that integrates the rules for its own organization. Learning does not mean acquiring information chaotically, but creating an order. In other words, knowledge forms a structure and not only a stock. To take Langlois and Garrouste’s words: ‘Knowledge is not a stock in the same sense that oil in a tank is a stock, something modified in a purely quantitatively way by the inflow or outflow of info-fluid. [...] Knowledge is about structure’ (Langlois and Garrouste 1997: 288). Therefore, to learn means both to reorganize the structure of existing knowledge, and to acquire some new knowledge. So there is a difference in kind between what is *received* (the stimulus), and what is *understood*, that is between the message, which comes from outside and the knowledge, which is a subjective and partial organization of it. As Nooteboom (2005: 37) points out, ‘information is not the same as knowledge [...] To become knowledge, information needs to be interpreted and understood in a cognitive framework’.

The second question that arises from Varela’s approach deals with the role of the individual and his subjectivity in the way he learns. As we showed before, the Simonian cognitivist approach assumes that thinking beings (whether they are human or ant), will always behave in a similar way, trying to adapt themselves to their environment. The idea on which is based the enaction

---

<sup>9</sup> ‘Like colors, smells are not revealed through a passive function of external properties, but through the creative articulation of a meaning that comes from history’ (Varela 1997: 111).

approach is on the contrary that the world that is perceived is not independent from whom perceives it. Here is what Varela writes in the preface of the second edition of his book:

‘One of the lead of this book is that ordinary life necessarily implies *embodied* agents, who are continuously confronted to the problem of action, whereas their different sensory-motor systems are engaged in parallel activities. [...] The corporality implies that the cognitive entity has – by definition – a perspective. It means that its links with the environment are not “objective”, independent from the situation, from attitudes and from the system history. On the contrary, these links are closely related to the perspective established by properties that keep emerging from the agent itself, and to the role played by these redefining in the coherence of the whole system’ (Varela 1997: iii).

In other words, the adapting process works first in relation with an *intention* (which he calls the ‘perspective’ of the cognitive system) that finds its origin in the many dimensions of the individual himself, in his history, his mental universe, his ‘corporality’, or in what we could more simply call his *identity*.

#### *A new relation between learning and rationality*

Bertrand Russel gave in 1954 the following definition of reason: 'It signifies the choice of the right means to an end that you wish to achieve. It has nothing whatever to do with the choice of ends' (Russell 1954: 8). Russel’s comment aims to define rationality in order to favour the relation between means and ends, and to keep it from a too normative aspect. Indeed, for Russel, the rationality criterion does not refer to a particular behavior, but to a logic of coherence between the behavior and the agent’s own finalities. It has to be said that this conception of rationality is not contradictory with the neoclassical or Simonian conceptions, in which there are adequacy between means and ends (even if the former is based on optimization while the second restricts itself to the ‘satisficing’ criterion).

Let’s now consider learning. In Simon’s view, as well as in the neoclassical model, learning is ‘substantive’. There is only one way to learn, the one that consists of memorizing true information about the world properties. But in the enaction approach, many different ways of learning are possible. The way you learn, the kind of data you memorize, depends on the perspective you choose. Learning becomes a creative activity, whose norms and frames are guided by the specific intention of the system. Paraphrasing Russel, a rational learning may then be defined in the following way:

*Rational learning signifies the choice of the right knowledge accumulation to an intention you wish to adopt. It has nothing whatever to do with the choice of intentions.*

In other words, being rational in learning does not mean accumulating objective information data in a universal logical process, but deepening a specific intention that may be subjectively determined.

Simon showed that two rational human beings will not necessarily behave in the same way, even if their objectives are identical, because they do not have the same view of the world. He implicitly assumed, however, that the way they learn would be similar. We therefore suggest to enlarge Simon's initial intuition about behavior, considering that, even in learning, the rationality criterion does not refer to an essence of learning, but to a principle of coherence between what is learned and what was intended to be deepened.

#### *Identity as the driving force of learning*

Learning processes are based both on equilibrium and on evolution principles (Cayla 2008). If changes in the cognitive system are necessary for learning, accumulation is also one of its main characteristics. This accumulation implies that a learning process has to maintain the stability of the knowledge that was previously acquired.

In the cognitivist approach, knowledge is seen as a stock of memorized information. As a consequence, learning is obviously cumulative, since the volume of information increases as the system perceives new environmental feedbacks. In the enaction approach however, it is the system's intentionality that gives a sense to external stimuli. Any change in this intentionality may therefore call into question the acquired knowledge; it is therefore important to preserve it in such a way that the learning process can go on.

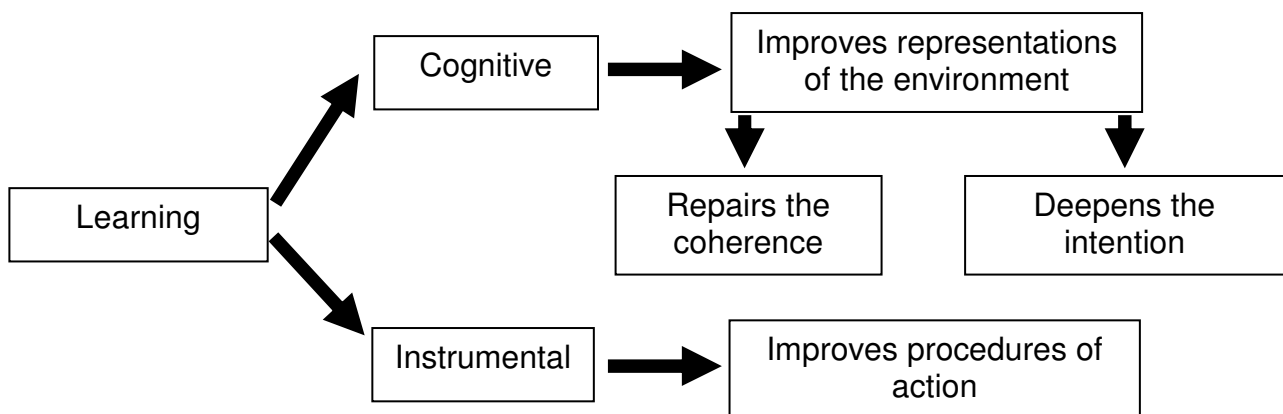
#### *Learning and coherence*

Unlike Simon's argument, the learning processes have to be considered, not as a mere mechanism that adapts itself to the environmental conditions, but as internal and complex processes that are continuously torn between their needs for internal coherence and the necessity for them to be flexible. This approach implies changing the angle of the analysis in a way to focus not only on environmental signals, but also on the internal mechanisms of

response (Marengo 1995). Therefore, cognitive learning processes have to deal with two antonymic constraints. On the one hand they have to let the organism develop a knowledge system that is appropriate to its environment; on the other hand, they must not affect the coherence that both maintains the acquired knowledge and allows accumulating some new learning. Dibiaggio (1999: 119) emphasizes that learning may take the shape of a 'coherence repairing process' that restores the cognitive system. Such a process implies an introspective work of learning that aims to realign the nature of acquired knowledge with the specific intentionality of the individual.

It has to be said that the adapting processes and the coherence repairing processes are two complementary kinds of learning. If coherence is not maintained, the organism is unable to assimilate and to give a sense to new knowledge. So the adapting process of the cognitive system should not be considered as an objective learning about the environment, but as a *deepening* process of the agent's intention about its environment.

The figure 2 bellow completes our representation of learning forms.



*Fig.: Extended representation of the different kinds of learning*

### *The notion of identity*

The last question we have to answer deals with the way the perspective of the cognitive system evolves. As it was said, the organism learns from an intention that helps it to acquire knowledge in a coherent way. But this intention may also change, according to an internal evolution, or in order to adapt the organism's knowledge to a change in its environment.

A way to conceive this change is to study the relation between behavior, knowledge and identity. The identity concept has been studied in numerous works in economy (Akerlof and Kranton 2000, 2005; John Davis, 2003). The French *Revue de philosophie économique* devoted

one of its special issues to this subject, among whom Kirman and Teschle, Alban Bouvier, Pierre Livet, John Davis, Amartya Sen, and Philip Grill contributed.

As Akerlof and Kranton explain, the identity notion can be very important in economy: 'Because identity is fundamental to behavior, choice of identity may be the most important 'economic' decision people make' (Akerlof and Kranton 2000: 717). Yet, in order to be 'economic', identity must be influenced by economic decisions, e.g. by conscious actions. The question is then to know whether agents have or not the freedom to 'choose' their identity. Sen's answer is that the notion of identity is in contradiction with the choice axiom, which is fundamental in economy: 'We 'discover' rather than choose our identity' (Sen 2004: 10).

Another question raised by the notion of identity is that it calls into question the assumption of the stability of preferences. If the identity of an individual changes, one may expect this individual to re-evaluate the utility of his actions and therefore his preferences. This approach takes us away from the homo economicus model, and implies that economists must study more closely the social and psychological roots of human behavior (Kirman and Teschl 2004, Sen 2004, Akerlof and Kranton 2005).

At last, accepting that the identity of an individual may change according to psychological or social criterions may endanger the very notion of 'identity'. How can an evolving person stay 'identical' to himself? For Davis, it is precisely the notion of identity that helps to comprehend the 'personal continuously' of an individual, the fact that an individual can stay 'the same' while he changes at the same time. This view is also adopted by Kirman and Teschl, who consider that the notion of identity goes beyond the question of preferences. Preferences have not to be considered as exogenous and unchanging, but must be seen as progressing within the framework of a constructing identity, so that individuals 'follow a certain path' (Kirman and Teschl 2004).<sup>10</sup>

If we adopt this view of identity, the way an individual changes its cognitive perspective and revises his intentions is internalized. The relationship between identity and intentionality can be specified following Grill's study of the concept of identity. Grill rejects both the essentialist and the existentialist views of identity and suggests adopting a 'pragmatic' perspective. He then defines identity as the capacity for a person to self-designate. But this self-designation is

---

<sup>10</sup> 'We might consider homo economicus as an agent *who* wants to do things and *who* is able to influence the world. The economic agent creates, builds, changes, and learns, is self-reflexive and evaluates her actions. She consciously affects and changes not only the conditions of life but also her own self-perception, thus her personal features and attitudes' (Kirman et Teschl 2004: 62).

not given once and for all, and is not independent from the social environment. It is constructed in a process that helps an individual to give a sense to his own actions, trying to ‘gather its life under the form of a story’ (Grill 2004: 90). So identity is not predetermined. It is not discovered as Sen argued, but it is ‘conquered’ (*Ibid*: 94). It emerges like a ‘life’s work’, through the story the individual gives himself.

This view of the concept of identity supports our analysis. Grill notably emphasizes that ‘life cannot be reified to a mere collection of behaviors and, to have a meaning, each action must be articulated with the life of the one who undertakes it. Actions may not be analyzed separately. They must be analyzed in relation to their coherence, in the person’s process of self-constitution’ (*Ibid*: 95). In other words, if intentionality is defined as the relationship that an individual develops with its environment and on the basis of which he constructs his cognitive system, then *identity can be defined as the relationship that an individual develops with himself and on the basis of which he creates his own intentionality*. As the existence of an intention allows the individual to interpret and comprehend its sensorial stimuli, ‘by telling himself a story and by interpreting his experiences, the individual discovers and reveals his own identity’ (*Ibid*: 96).

We can now develop a global analysis of the learning process, by distinguishing three overlapping and hierarchical levels that may help to comprehend the working of a cognitive system (figure 3).

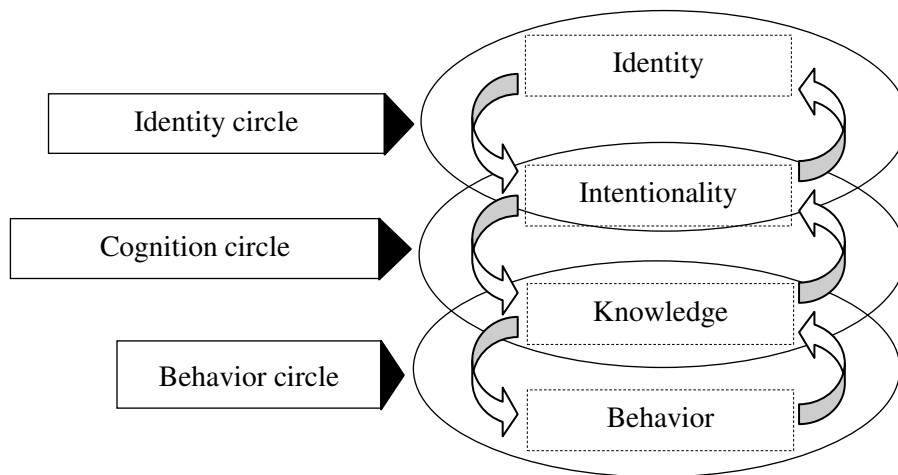


Fig. 3: Representation of the cognitive system

The lesser level, the ‘behavior circle’, aims to develop the relationship between the behavior and the system of knowledge. Thus, behavior is determined by the level of knowledge, but this knowledge itself is constructed according to a desired or future action. We find again the

Varela's view of enaction, which postulates that action and cognition mutually defined themselves.

One level above, the 'cognition circle', represents the system's cognitive learning. This type of learning is connected with an intention, which helps to organize the knowledge as a coherent and individualized structure instead of a stock of objective information data. Here again, knowledge and intentionality codetermine themselves. Knowledge is partially the product of intention, whereas intentionality is also a construction that is based on the individual's knowledge and on the way he looks at his own actions.

At last, the third level, the 'identity circle' expresses the deeper learning level, the one that helps the individual to 'see' himself. With the construction of his identity, the individual builds his intentions, and it is with his intentions and through the way he sees himself, that his identity is built.

This representation clearly assumes two fundamental ideas. First, it is based on the assumption that there are no frozen level and that the whole cognitive organization of an individual may change. Second, it shows the hierarchical aspect of this organization, implying that a change in one of these levels does not involve the same kind of learning in any case. A change in the relationship between knowledge and behavior (an instrumental level) will less destabilize the cognitive structure than a change in the relationship between intentionality and identity. In order to maintain its cumulative dimension, a learning process should therefore make a frequent use of low-level learning process, and should keep certain stability in the architecture of the cognitive system, given by its identity. The question is then to understand the trade-off between what an individual has to 'unlearn' in order to implement a new behavior in a given environment.<sup>11</sup>

## 5. Conclusion

Learning means two quite different things. For most economists, it means to bring knowledge progressively closer to a given reality. Thanks to learning, I can acquire information that completes and improves my representation system, which tends to be a mirror that will be smoother and smoother, and that will better and better reflect my environment. Although this approach of learning has the advantage to be compatible with the neoclassical and Simonian

---

<sup>11</sup> It is worth noting that these three cognitive levels are related to Bateson's (1972) three levels of learning. Interestingly, Bateson notices that learning III involves a 'deep reorganization of one's character' and that it may call into question the individual's definition of himself.

rationality models, it seems inadequate to comprehend learning phenomena for at least three reasons:

- It cannot conceive learning as a process, *i.e.* an ordered succession of steps that has its own logic. Instead, learning is simply reduced to a chaotic information data accumulation.
- It cannot comprehend learning as the product of a *choice*. Learning comes down to a linear improvement of the performances of the relationship between the individual and his environment.
- The problems related to the convergence or to the coordination of learning by different persons are assumed to be non-existent. When different agents are confronted to the same environment, they are supposed to learn the same thing. As a consequence, their cognitive diversity is supposed to decrease as their knowledge of the world increases. This assumption is especially problematic if one wants to study collective or organisational learning, in which coordination problems are the core of the analysis.

In this paper, we chose to try another approach that defines learning as the guided development of a cognitive system, instead of as an accumulation of ‘true’ information. In our view, learning does not bring knowledge closer to an objective reality, but aims to deepen the progression of a cognitive system in its own reality. This implies that there is no ‘pure’ and unequivocal information. Thus, if I learn, I will not construct a mirror that will reflect the environment with more and more faithfulness, but I will establish a partial and subjective view of it.

This second approach of learning solves the difficulties of the first one but has two important consequences.

- First, learning of two different individuals will not make them converge toward the same representation system, and to learn will not help them to understand each other better. Learning processes may on the contrary increase their *divergence*, because they will not learn the same thing from the same stimulus. As a consequence, individual learning may destabilize a social organization.
- The second consequence of this approach deals with the rationality model that has to be adopted. In this paper, we showed that there is not only one rational learning model, but many of them, for what matters in learning is the level of *coherence* between the learning process and the individual’s intentionality and identity. Learning in this meaning resemble the product of an internal and complex system that is continuously

torn between its needs for internal coherence and the necessity to adapt to its environment.

## References

- AKERLOF, George A. and Rachel E. KRANTON (2000), "Economics and Identity", *The Quarterly Journal of Economics*, vol. 115, 3 p. 715-753.
- (2005), "Identity and the economics of Organizations", *Journal of Economic Perspectives*, vol. 18, 1, p. 9-32.
- ARENA, Richard and Nathalie LAZARIC (2004), "La théorie évolutionniste du changement économique de Nelson et Winter", *Revue Économique*, pp. 329-354.
- BATESON, Gregory (1972), *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. University Of Chicago Press.
- BESSY, Christian (2003), "Distributed Cognition and the Theory of the Firm: Enriching Nelson and Winter's Evolutionary Analysis of Organizational Learning and Memory", *Économie Appliquée*, vol. 56, 4, p. 65-120.
- BINMORE, Kenneth (1987), "Modeling Rational Players", *Economics and Philosophy*, vol. 3, p. 9-55.
- BOUVIER, Alban (2004), "Choix rationnel et identité. Le cas des conflits nationalistes et ethniques", *Revue de Philosophie économique*, 9, p. 135-160.
- BÖRGERS, Tilman, (1996), "On the Relevance of Learning and Evolution to Economic Theory", University College London, mimeo.
- CAYLA, David (2008), "Organizational Learning: A process between Equilibrium and Evolution", *Journal of Economic Issues*, Vol. 42, 2, p. 553-559.
- DAVIS, John B. (2003), *The theory of the Individual in Economics*, London: Routledge.
- (2004), "Complex Economic Systems: Using Collective Intentionality Analysis to Explain Individual Identity in Networks", *Revue de Philosophie économique*, 9.
- DIBIAGGIO, Ludovic (1999), "Apprentissage, coordination et organisation de l'industrie une perspective cognitive", *Revue d'Économie Industrielle*, 88, p. 111-136.
- DOSI, Giovanni, Luigi MARENGO and Giorgio FAGIOLO (2003), "Learning in Evolutionary Environments", LEM Working Paper Series.
- ELLISON, G. et FUDENBERG, D. (1993), "Rules of Thumb for Social Learning", *Journal of Political Economy*, vol. 101, p. 612-643.
- FOSS, Nicolai, J. (2003), "Bounded rationality and tacit knowledge in the organizational capabilities approach: an assessment and a re-evaluation", *Industrial and Corporate Change*, vol. 12, 2, p. 185-201.
- GARROUSTE, Pierre (1999), "Apprentissage, interactions et création de connaissance", *Revue d'Économie Industrielle*, 88, p. 137-151.
- (2003), "Learning in Economics: Some Austrian Insights", in Rizzello, S. (ed.) *Cognitive Developments in Economics*, London: Routledge, p. 302-315.

- GRILL, Philippe (2004), “Une approche intrinsèque de l’identité pour les sciences sociales”, *Revue de Philosophie économique*, 9, pp. 87-114.
- KIRMAN, Alan and Mark SALMON (1995), “Editors Introduction”, in *Learning and Rationality in Economics*, pp. 1-11, Oxford: Blackwell.
- KIRMAN, Alan and Miriam TESCHL (2004), “On the Emergence of Economic Identity”, *Revue de Philosophie Economique*, 9, p. 59-86
- LANGLOIS, Richard N. and Pierre GARROUSTE (1997), “Cognition, Redundancy, and Learning in Organizations”, *Economics of Innovations and New Technology*, vol. 4, p. 287-299.
- LAZARIC, Nathalie, Jean-Michel MONNIER and Bernard PAULRE (1995), “Introduction Générale”, in Lazaric, Monnier et Paulré (eds), *Coordination économique et apprentissage des firmes*, Paris: Economica, p. vii-xx.
- LIVET, Pierre (2004), “La pluralité cohérente des notions d’identité personnelle”, *Revue de philosophie économique*, 9, p. 29-57.
- MARENGO Luigi (1995), “Apprentissage, compétences et coordination dans les organisations”, in LAZARIC N and MONIER J M (eds.), *Coordination économique et apprentissage des firmes*, Paris: Economica, pp. 3-22.
- METCALFE, J. Stanley (2005), “Evolutionary concepts in relation to evolutionary economics”, in *The Evolutionary Foundation of Economics*, Cambridge, MA: Cambridge University Press, p. 391-430.
- NELSON, Richard R. and Sidney G. WINTER (1982), *An Evolutionary Theory of Economic Change*, Cambridge, MA, and London: The Belknap Press of Harvard University Press.
- NOOTEBOOM, Bart (2005), “A cognitive Theory of the Firm”, *European Journal of Economic and Social Systems*, vol. 18, p. 29-59.
- PAULRÉ, Bernard (2005), “Towards a Connexionist Theory of the Firm”, *European Journal of Economics and Social Systems*, vol. 18, p. 13-27.
- RUBINSTEIN, Ariel (1998), *Modeling Bounded Rationality*, Cambridge MA: The MIT Press.
- RUSSEL, Bertrand (1954), *Human Society in Ethics and Politics*, London: George Allen.
- SEARLE, John (1983), *Intentionality: An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press.
- SEN, Amartya (2004), “Social identity”, *Revue de Philosophie Economique*, 9, p.7-27.
- SIMON, Herbert, A. (1986), “Rationality in Psychology and Economics”, *The Journal of Business*, Vol. 59, p. 209-224.
- (1996), *The sciences of the artificial*, 3rd ed. MIT Press.
- (2005), “Darwinism, altruism and economics”, in Kurt Dopfer (ed.), *The Evolutionary Foundation of Economics*, p. 89-104, Cambridge: Cambridge University Press.
- VARELA, Francisco, J. (1997), *Invitation aux sciences cognitives*, Paris, Éditions du Seuil.
- WALLISER, Bernard (1995), “Rationalité instrumentale et rationalité cognitive”, in L.A. Gérard-Varet and J.-C. Passeron (eds.), *Le modèle et l’enquête, les usages du principe de rationalité dans les sciences sociales*, Presses du CNRS.
- (1998), “A spectrum of equilibration processes in game theory”, *Journal of Evolutionary Economics*, vol. 8, p. 67-87.
- (2000), *L’économie cognitive*, Paris: Odile Jacob.