



University of Navarra

Working Paper

WP no 654

October, 2006

BEYOND ECONOMIC CRITERIA: A HUMANISTIC APPROACH TO ORGANIZATIONAL SURVIVAL

Josep M. Rosanas

IESE Business School – University of Navarra

Avda. Pearson, 21 – 08034 Barcelona, Spain. Tel.: (+34) 93 253 42 00 Fax: (+34) 93 253 43 43

Camino del Cerro del Águila, 3 (Ctra. de Castilla, km 5,180) – 28023 Madrid, Spain. Tel.: (+34) 91 357 08 09 Fax: (+34) 91 357 29 13

Copyright © 2006 IESE Business School.

BEYOND ECONOMIC CRITERIA: A HUMANISTIC APPROACH TO ORGANIZATIONAL SURVIVAL

Josep M. Rosanas

Abstract

There are many theories about organizations that are mutually inconsistent with each other, which explain phenomena to very similar extents. Most of them ignore the ethical dimension completely. In this paper I put forth the basic principles for a theory of decision-making in organizations which integrates ethics in the core of the theory. It is based on the work of Juan Antonio Pérez López (1991, 1993) and is essentially a humanistic view of the interrelationships between people and its implications for organizational decision-making. I will first show that in any relationship between two people, the learning of the two is crucial for such a relationship to last; and then I will expand on the different aspects of that learning. This analysis will then be applied to the organizational context as a basis for organizational decision-making. Second, it applies the previous analysis to the organizational context as a basis for organizational decision making, showing how any decision in an organization needs to be analyzed on the basis of three criteria (short-run effectiveness, development of distinctive competence, and unity and identification with the organization) and how ethics is included in the last two.

Keywords: business ethics, ethical foundations of organization, trust, management theory, learning.

BEYOND ECONOMIC CRITERIA: A HUMANISTIC APPROACH TO ORGANIZATIONAL SURVIVAL

Introduction¹

We could possibly find wide agreement in the elementary fact that organizations exist in order to coordinate different individuals' efforts towards achieving some objective that a single person alone would be unable to achieve. But beyond that, if we examine the different organizational theories that have had wide diffusion in the world of management, we won't be able to find too much of an agreement in anything else. Organization theory has greatly evolved throughout the years and different schools of thought have had substantially different views of what theory is and should do. Ghoshal (2005, p. 86) goes as far as stating that "many different and mutually inconsistent theories explain the same phenomenon, often to very similar extents".

A classic book on organization theory (Scott, 1987) classifies the different schools of thought in three groups: rational, natural and open systems. Scott's definitions of each one clearly show the wide gap between them, although, as Scott says, there have also been some attempts to integrate the three perspectives.

Most organizational theories don't explicitly consider two crucial factors: learning and ethics. In this paper, I present the basis of a theory of organizations that explicitly includes them, and identifies the link by which they are related. I will proceed as follows. First, I will show how two of the most used and cited approaches in organization theory (agency theory and institutional theory) do not include learning and ethics in their framework. I will then attempt to establish the bases of a different theory that takes learning into account, and show that there are different aspects of learning that have strong implications for ethical considerations. These bases are established by examining the dyadic relationship between two people in some detail, and the starting point is the (boundedly) rational behavior of the agents. Finally, by using the

¹ A previous version of this paper was originally presented at the IESE Business School, University of Navarra, for the 14th International Symposium on Ethics, Business and Society: "Towards a Comprehensive Integration of Ethics Into Management: Problems and Prospects". May 18-19, 2006). I am grateful to Rafael Andreu and Domènec Melé, for valuable and detailed comments on earlier versions of the paper.

previous model, I will present some conclusions on how decision-making should take place in organizations to make them survive in the long run.

Theories of management

Most of the theories about organizations differ in several dimensions. One of them (only implicit in the Scott formulation mentioned above) is the degree of formalization that they have, which is to some extent related to their degree of rationality. The economic approaches (economics being considered the realm of rationality) are in general quite formalized, although some of them are much more formalized than others. Agency theory, to name one of the most popular, is as we will see, one of the most formalized (at least in one of its versions); while theories that in Scott's typology fall under the heading of 'open systems' show a much lower degree of formalization. Another dimension is realism: agency theory has been criticized (Donaldson, 2002; Ghoshal, 2005; Ferraro, Pfeffer and Sutton, 2005) for making unrealistic and pessimistic assumptions about people's behavior in organizations, while institutional theory is supposedly based on empirical observations of the real world.

Agency theory

As suggested above, agency theory comes at least in two versions, substantially different in many respects (Jensen, 1983; Nilakant and Rao, 1994): the so-called 'principal-agent theory', which is the more formalized version already mentioned, stems from the work of Ross (1973), Holmström (1979), and Grossman and Hart (1983); while the so-called 'positive' agency theory, that stems mainly from the work of Jensen and Meckling (1976). The first one is thus highly mathematical, very precise in its assumptions and very rigorous in its deductions. The second one, the 'positive' version, is based on verbal and graphical analyses, (purportedly) more empirical (Jensen, 1982) and prone to practical recommendations. It is this second one that is more often subject of the criticisms that have been mentioned, the first one being far more rigorous, clearer about its foundations and willing to incorporate more realism in their basic assumptions if they only can be treated with the same rigor.

Both versions start from the basic economics assumptions of the *homo oeconomicus*. Under those assumptions, human beings are not 'real people', but 'utility maps' instead, which reflect unambiguously and with unbounded rationality their wants and preferences, typically based on self-interest and essentially between material goods; and they try to satisfy those wants and preferences. This is more explicit and formal in the principal-agent literature, where it makes it easier to see when the theory attempts to go beyond the conventional 'homo oeconomicus' to a different species of 'homo' with wider interests.

The 'principal-agent' literature starts also from a hierarchical hypothesis: that the principal is 'in command' and is the person whose welfare should be maximized, while the agent should execute the orders and/or wishes of the principal by receiving an 'adequate' compensation, related perhaps with the agent's opportunity cost 'reservation utility', in the language of the theory; and which should be as small as possible in order to make the principal's welfare as big as possible. Efficiency, in the Pareto sense, is the main goal of the theory. This goal excludes inefficient possibilities that might benefit the principal at the expense of the agent, but the whole theory is biased towards the principal in the sense that it is the principal who designs the compensation system of the agent, and thus, has the upper hand in choosing between different efficient combinations.

The arguments of both utility functions (the principal's and the agent's) are essentially monetary values; although from rather early, some principal-agent models have included disutility for effort on the part of the agent as well (i.e., an 'effort' variable of which utility is a decreasing function, and 'results' an increasing function). Later, more sophisticated models have attempted to include other variables, but the extent to which they have achieved it within the formal model is rather limited. Variables like 'identification', 'loyalty', which are considered in 'rational' approaches like that of Simon (1947), or 'caring for other people's welfare' are unknown to the agency literature.

Human beings interpreted as utility functions can be considered a 'mechanical' model, but at the same time, their utility represents their intentions. Thus, agency theory (like the rest of organizational economics) can be seen as an attempt to produce at the same time what Elster calls 'causal' and 'intentional' explanations (Elster, 1983). In fact, unbounded rationality, paradoxically perhaps, makes human beings so intentionally calculative (taking into account the very long run as easily as the short run) that their behavior becomes perfectly foreseeable, or 'mechanical' and, thus, explanations may become 'causal'. Bounded rationality, in contrast, makes causal and intentional explanations very different.

Institutional Theory

If we go now to institutional theory, we can see how things look quite different from there. As in the case of agency theory, institutional theory has at least two substantially different, almost opposite, versions as well: the 'old' and the 'new' institutionalism (Selznick, 1996). This paper will take the 'modern', or 'new' version of institutionalism as a reference. According to the path-breaking authors in the area:

"The new institutionalism in organization theory and sociology comprises a rejection of rational-actor models, an interest in institutions as independent variables, a turn toward cognitive and cultural explanations, and an interest in properties of supraindividual units of analysis that cannot be reduced to aggregations of direct consequences of individuals' attributes or motives." (DiMaggio and Powell, 1991, p. 8). And also: "Organizations tend to model themselves after similar organizations in their field that they perceive to be more legitimate or successful. The ubiquity of certain kinds of structural arrangements can more likely be credited to the universality of mimetic processes than to any concrete evidence that the adopted models enhance efficiency." (DiMaggio and Powell, 1991, p. 70).

Selznick comments on mimetic processes that "Mimesis is considered a response to uncertainty presumably more deeply rooted in anxiety than in rational efforts to avoid reinventing the wheel. This suggests, perhaps, that organizational adaptation is often more compulsive than problem solving" (1996, p. 273). Donaldson (2002, p. 101) provides a rather critical summary of institutional theory:

"Institutional theory argues that organizations conform to norms held in their environments about sound organizations. Conformity comes from taking things for granted, adherence to norms, coercive sanctions, or the desire to gain legitimacy and resources from external organizations and professions (Powell and DiMaggio, 1991). The implication is that the managers running organizations are following unconscious 'taken for granted' or seeking certainty in the face of causal ambiguity or, more calculatively, trying to make their organization look good to an external audience. (...) The bold claim of institutional theory is that organization is not about rationality, but about ritual and conformity to some ideology, such as the ideal of bureaucratic rationality (Meyer and

Rowan, 1977). The stress on ritual over rationality runs the risk of undermining the educational mission of universities to promote reason in their students and in their alumni who are managers.”

Thus, institutional theory assumes a process of isomorphism by which organizations come to resemble other organizations; and the practical application of such a principle is that firms should adopt the same procedures and decisions as the competitors; which is something that goes directly against the idea of a differentiated strategy that is generally accepted to be the key to higher profitability.

But criticisms aside, what is rather obvious is the incompatibility on several accounts between institutional theory and agency theory, which might be considered the two extremes of the theories of organization. First, institutional theory is not structured in formal terms, which means that the basic assumptions of the theory are more implicit than explicit. Second, the idea of legitimacy and ritual is at the antipodes of a clear preference ordering or intentionality. Third, the kind of explanation that such a theory provides on organizational phenomena is more functional than causal or intentional. Finally, what human beings value in this theory is far from clear.

The two theories are, thus, not only incompatible, but also incomparable to a great extent, given the distance between their starting points and their perspectives. From a different point of view, perhaps closer to agency theory, but also incompatible with it, we can find stewardship theory (Davis, Schoorman, and Donaldson, 1997), cited by Ghoshal (2005) as a “much more sensible” possible alternative that takes into account the interests of many people simultaneously (customers, employees, shareholders, communities where they operate...). The approach presented below may be considered a way to further develop the stewardship theory in a more formal way and approach it to a different form of agency theory.

Learning and Organization Theory

Interestingly, one similarity between the two theories briefly examined above is that learning plays in them a rather small role, if any. Only in some extensions of agency theory (more related with game theory than with agency theory itself) a very limited concept of learning is considered, that takes into account basically the updating of the agents’ beliefs both about nature and about the other agent, as we will see immediately. Institutional theory has very little place for learning as well. While one might argue that ‘mimesis’ to gain legitimacy is related with learning, because it changes people’s behavior, this concept of learning is rather limited (not related with ‘reasoning’ at all) and does not fit with the intuitive idea most of us have about what learning is.

Since the main objective of this paper is to start from a (boundedly) rational approach to organizations to expand its basic assumptions in the direction of making it more realistic, it becomes important now to go deeper into the concept of learning in the economics-based theories to show what it includes and what it does not. For that purpose, the cases of unbounded and bounded rationality will be examined according to the classical concepts developed by Herbert Simon (1947).

Unbounded Rationality

In most economic theories, but especially in agency theory, the assumption of unbounded rationality strongly conditions the conceivable concepts of learning applicable to any situation. In effect, with unbounded rationality: (i) agents possess some knowledge about the real world, (or 'nature', as it is often called in the theory of decision-making under uncertainty), which is represented by a probability distribution of the possible states of nature before and after any given action is taken, and (ii) agents have an unambiguous utility function that perfectly represents their preferences. As is well known, unbounded rationality does not imply certainty, which is why the two previous points take uncertainty into account: agents' knowledge of the real world includes a probability distribution about the 'state of nature' and their utility function includes preferences for uncertain outcomes.

Thus, when an unboundedly rational agent interacts with 'nature' (assuming for a moment no interactions with other individuals), the results make agents 'learn': they change (update) their probability distribution according to the laws of probability, specifically, the Bayes Theorem (bayesian learning)². If, besides, they interact with another agent, and since under unbounded rationality all agents have a clear and definite preference ordering, then they each belong to a 'class' of individuals with that preference ordering. For each single agent, 'knowing' the other agents means having a probability distribution about the 'class' to which they belong; and 'learning' about them, then, reduces to the updating of such beliefs (again, according to the Bayes Theorem). In summary, with unboundedly rational agents, the concept of learning is merely calculative, and merely consists of an updating of probabilities about nature and about the other individual(s).

Bounded Rationality

Bounded rationality changes the picture substantially. First, with bounded rationality, there is Bayesian learning as well, but subject to 'mistakes', that occasionally result in 'superstitious' learning, attributing the wrong causes to observed results, or 'confirming' one's own biases (March and Olsen, 1975; Yariv 2002). Bayesian updating includes considering unexpected results that were assigned zero probability before an interaction, i.e., increasing the agent's ability to foresee the unforeseen.

Second, agents may learn by increasing their ability to put into practice the same action, i.e., doing it with less effort, better results, or both. External, explicit knowledge does not change, but the efficiency in its use does. This is the meaning of learning behind the familiar 'learning curve' (or, more in generally, the 'experience curve') that intends to represent the decrease in unit costs that occurs when there is a cumulative increase in the number of units produced.

Third, with bounded rationality agents are not too sure about their preferences, and often realize how much they like (or dislike) the results of a given action or the action itself (effort, etc.) only when it is already done and they can evaluate the results, which may not be available immediately after the action, because some results take place only in the long run. This is closely related to what has been analyzed in the economics literature under the heading of 'time inconsistency of preferences'. In the words of O'Donoghue and Rabin: "Casual observation, introspection, and psychological research all suggest that the assumption of time-consistency is

² For an explanation of how this updating process takes place, see, for instance Marschak and Radner (1972).

importantly wrong. It ignores the human tendency to grab immediate rewards and to avoid immediate costs in a way that our ‘long-run selves’ do not appreciate” (1999, p. 103)

Learning whether the (short term and long term) results are worth the effort and the actions taken is a type of learning that is completely different from the Bayesian learning analyzed above. Following Pérez López, we will call it ‘evaluative learning’ (1993, mainly chapters 10 to 12).

There is also another aspect of evaluative learning that will be discussed below. It has to do with the attitudes agents have with respect to each other: one agent may ‘like’ and/or ‘trust’ the other or not. Trust between agents is developed through interactions that allow them to gain a better idea of each other’s value system and decision-making characteristics (Rosanas, 2004).

The theory to be introduced in this paper attempts to accommodate the different kinds of learning that organizational actions induce in the people involved as the cornerstone of organizational analysis, as will be discussed below. All these concepts about learning are summarized in Table I.

Table 1. Decision-making and learning in interactions

	Unbounded rationality	Bounded rationality
Interactions with ‘nature’	-Bayesian learning (updating beliefs according to prior and data from experience).	-Bayesian learning (perhaps incorrect or superstitious learning - biases). Includes foreseeing the unforeseen. -Acquiring or improving skills. -Predicting actual ‘after the fact’ utility (evaluative learning).
Interactions with another human being	-Bayesian learning of both agents about Nature. -Bayesian learning of both agents about each other’s type.	-Same as in interactions with nature for each agent (i.e., Bayesian learning, unforeseen consequences, skills, evaluative learning). -‘Trust’ in the other agent. Possibility of ‘negative’ learning.

Ethics and Organization Theory

Most management theories ignore ethics almost completely. The two we have taken as points of reference, agency theory and institutional theory, are good examples of this. as stated, agency theory emphasizes Pareto-efficiency without taking into account ethical consideration. The unboundedly rational *homo oeconomicus* has no ethical doubts or problems: since he knows his preferences and wants perfectly, can act in accordance with them, and those preferences are completely arbitrary, anything he prefers or does in practice can be thought to be ‘ethical’. Agency theory, if anything, implicitly takes sides with principals, but the main point is that it often ignores issues such as lying, cheating, abuse of power, and so on; or, even worse, it takes

them for granted (the principal will never know the private information about his/her effort that the agent possesses, and therefore the principal will never know the truth and the agent can lie about it without restraint). At best, when it does not take them for granted, there is an implicit assumption that this does not happen, or is not a problem, while in management practice this is often *the* problem.

Institutional theory does not (and possibly cannot) even address ethical considerations, because ethics is based on rationality, as this paper will argue, and the 'raison d'être' of institutional theory is precisely non-rational behavior. Donaldson's criticism of the theory quoted above (2003) is based on its lack of rationality ("stress on ritual over rationality"). The morality of mimetic behavior is not even considered: it is only its 'legitimacy' that counts. Social values are taken as given and accepted, leading to a concept of relativistic ethics that is quite problematic (Frederick, 2002) Where economic models accept any arbitrary preferences by individuals, which may be of any kind, institutional theory blindly accepts the generally accepted social values.

In this regard, if economic models make pessimistic assumptions about human nature, the assumptions of institutional theory are perhaps even more pessimistic. While economic models assume that individuals are self-centered and perhaps even opportunistic, institutional theory assumes they cannot even be rational.³

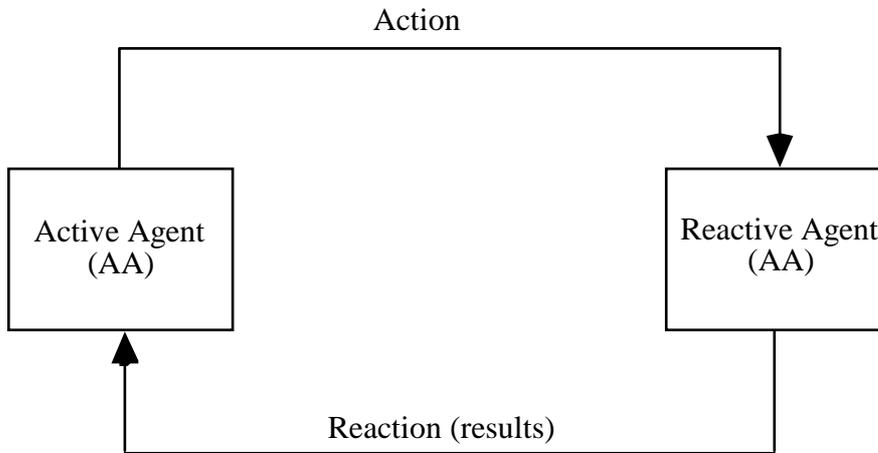
To integrate ethics and management theory we need a different approach. First, it has to provide intentional explanations and be realistic. Second, it has to promote rationality. And third, it has to include ethical concepts. What follows, is an attempt to put forward the elements of such an approach, which in fact has that integration built into the theory itself. It borrows heavily and freely from the basic concepts of the work of J.A Pérez López (1991, 1993), the essential elements of which were sketched by Rosanas and Velilla (2003) and Argandoña (2007).

A Dynamic, Symmetrical Model of Agency

The model proposed by Pérez López is based on the analysis of a situation somewhat similar to that of Agency Theory, although the concept of human nature and of human action used are quite different. The similarity is in the fact that it affects two individuals or economic agents, where one (the 'Active Agent', or 'AA') wishes to obtain some explicit results through the cooperation of another person (the 'Reactive Agent', or 'RA'). The Active Agent wants to take some "action" in order to obtain the desired "reaction" from the Reactive Agent (see Figure 1). The relationship between two people is thus the building block of organizational action both here and in Agency Theory.

³ This applies to the 'new' institutional theory. The 'old' one, that of Selznick (1957), for instance, was at the antipodes of that: "infusing organizations with value" was perhaps the main idea of the theory. See Selznick (1996) for a comparison of the two.

Figure 1



In contrast with agency theory, though, a non-hierarchical, symmetrical relationship between the two agents involved is assumed, where each one interacts with the other and they both have different (possibly private) interests to satisfy. There is, to be sure, an element of asymmetry in the relationship: one of the agents ‘starts’ the action, trying to solve a problem or achieve an objective, and the other one ‘reacts’, i.e., AA ‘starts’ the process and RA ‘follows suite’. But no assumption is made about any hierarchical relationship between AA and RA. Of course, a hierarchy may exist, but not necessarily. AA might be the ‘boss’ and RA the ‘subordinate’, but it might be the other way around, or they might both be at the same level in the organizational ladder. For instance, a production manager may want to obtain some explicit results in terms of a cost reduction, and may have to do it through the head of a given department. This is a hierarchical relationship in the same line of thought as in Agency Theory: the production manager is AA and the department head is RA. But this theory also considers the opposite situation: the department head of a department is interested in the company buying a new machine, and has to convince the production manager to agree with it and take the necessary steps with the management committee and financial officers for the purchase to be made. In this case, the department head is AA and the production manager is RA. In turn, the production manager may become the AA with respect to a financial manager at the same level, who then becomes the RA. More generally, in other contexts, AA and RA might just be ‘friends’: in fact, the proposed model of analyzing interactions can be used in a wide variety of settings.

A crucial element of the theory is the fact that, besides the external, explicit results, there are ‘internal’ consequences of the action–reaction dyad: the learning of the two agents. Such learning will (partly) determine their behavior in future interactions, and is thus quite important for the future of their relationship (and, as we will see below, to the survival of the organization if the interaction takes place in an organizational context). ‘Learning’ essentially means a change in the decision rule. We say that an agent has ‘learned’ taking an action at time t under a given situation, if there exists a situation where the agent would have responded with a

decision D_t , at time t , while at time $t+1$, facing a similar situation would respond with a different decision D_{t+1} , different from D_t ,⁴

Learning thus becomes a crucial element in the dynamics of interactions and is endogenous to them. If there were no learning, interactions in similar situations could only be repeated indefinitely. Therefore, the learning of both agents will partly determine the future, and it is now important for us to go back to the different types of learning that might take place and were briefly introduced above. As discussed there, bounded rationality implies the existence of substantially different kinds of learning. For our purposes here, it is important to distinguish between:

- a) *Operational learning*, i.e., the type of learning that increases the ability of agents to perform the corresponding action more efficiently. This includes, as stated above, improving their knowledge about the real world, the state variables, or the laws of nature, which can be summarized in a probability distribution about possible results that might occur following their actions. It essentially consists of explicit knowledge, relatively easy to transmit, although possibly difficult to authenticate under some circumstances. Every interaction provides data for (Bayesian) updating of the probabilistic information all agents have about nature. Operational learning also includes increasing the agents' ability to perform specific tasks (skills), making them more efficient to perform the corresponding actions.
- b) *Evaluative learning*, i.e., the type of learning that allows agents to evaluate a priori how much they will actually like the result of the action after the fact, or to what extent they think it fits their needs. This includes better knowledge of their own time preferences and their logical consistency; but it also includes possible changes in their attitude towards each other, the extent to which they 'trust', or have 'willingness to cooperate' with the other agent.

Under unbounded rationality, operational learning can only have 'positive' effects, i.e., effects that lead to better results for both agents, perhaps in terms of increased efficiency (increased results or decreased costs), perhaps in terms of decreased uncertainty. Under bounded rationality, in contrast, learning may be 'superstitious' (March and Olsen, 1975) i.e., agents may learn things that are not true and attribute the results to the wrong cause, thus leading in general to worse results in the future.

Evaluative learning can by definition have 'positive' or 'negative' effects, i.e., the agents may like the results better or worse than they thought, and the attitude of one agent towards the other may change in favor or against the other.

Of course, if there is superstitious learning and, besides, agents evaluate that the results had worse affects on their well-being than they expected and they trust less each other, then the results of the next interaction can be expected to be worse and the relationship between the two people will deteriorate even more.

⁴ There is an obvious difference between a single decision and a decision rule. Rigorously, in logico-mathematical terms, there may be specific decisions that do not change when the decision rule does. But if in no situation the decision changes, then it must be that the decision rule has not changed, or that the two are equivalent. Provided, then, there exists one single situation (as the text states) where the decision changes, there must have been a change in the decision rule, and, thus, learning.

Effectiveness and Learning

For AA, the effectiveness of his/her action depends on whether the expected results are achieved or not. The desire for such expected results is what starts the interaction in the first place; therefore, whether the action is effective or not is a crucial variable for AA. But obtaining these results does not depend only on him/herself, but (mainly perhaps) on RA as well. Suppose, for instance, that AA tries to use persuasion to convince RA to do something from which he/she expects some results. The results depend on AA's ability to persuade RA in the first place, but also on RA's ability and strength of will to perform the required tasks right.

'Persuading RA' essentially means convincing him/her to evaluate the proposal positively, i.e., that it is 'worthwhile', or in accordance with his/her own motives. Obviously, this would not be necessary in a world of unbounded rationality: RA would perfectly know beforehand his preference map, and therefore whether the proposal is in accordance with his/her motives or not as well (given current constraints); while in a world of bounded rationality this depends on the evaluative capacity of RA, who can make mistakes and change his/her mind (i.e., learn) through time. Notice that AA can 'tell the truth' to RA, showing him/her how the action proposed is in his/her own best interest, or try to 'fool' RA into believing that it is when it is not. In turn, RA may be able to discover whether the first or the second happened. Whatever actually happens will condition the evaluative capacity of RA for the future, both in abstract (i.e., in terms of the action itself and its results) and in relation with AA (i.e., whether AA 'can be trusted' or not). Thus, if RA feels that, in summary, the balance of action, results and the interaction with AA is unsatisfactory, it will be more difficult next time for AA to use persuasion to obtain the same results.

'Doing it right' depends on RA's ability and knowledge. A change in such ability and knowledge is what we have called 'operational learning'. Typically, operational learning will accumulate over time with practice, thereby increasing the capacity of the person to do things right. The exception is the already mentioned superstitious learning, when an agent learns something that is not true.

Therefore, the knowledge (of all kinds) that RA has is crucial for AA to obtain the desired results through an interaction; and learning includes the change in all that knowledge that will take place in the interaction, which will (partly) determine the effectiveness of future action-reaction dyads between the two agents.

Notice that neglecting to consider RA's learning is dealing with RA (a human being) as if he/she were a machine, and ignores therefore the basic premise of management, which consists of doing things through other people. When a human being, acting as an AA, interacts with a machine (the machine, in this case, being the 'RA'), that human being does not have to worry about the learning of the machine (which is simply inconceivable), but must only make sure that the machine is functioning properly, i.e., must know the current state of the machine. In contrast, when interacting with another human being, AA has to worry about the learning of RA, if only because the effectiveness to be expected in the future depends on the learning that takes place now. Thus, only a short-sighted AA (i.e., an AA not willing to look into the future) would not take RA's learning into account when making a decision for which he/she needs RA's cooperation.

AA's own learning is also important to him/herself for obvious reasons: it conditions the future for both AA and RA. It may be helpful at this time to illustrate it through an example. We said that because of bounded rationality, AA might obtain the desired results, and be unhappy with them; or may obtain results that are different from those expected and be happy with them. Suppose then that a General Manager decides to ask the Marketing Manager of the company to shoot for a 10% increase in sales. Suppose further that the Marketing Manager and the Marketing Department achieve it; and then they all see that the increase has had bad effects in costs, which have increased disproportionately because the increase in sales was achieved at the expense of constant production rescheduling. Probably, there has been operational learning in the process; and the General Manager may have (evaluatively) learned that increasing sales by 10% was not so desirable as he/she thought. Both types of learning are bound to have an impact on future goals proposed by the General Manager, and on his/her interactions with the Marketing Manager.

Finally, as we mentioned above, creating trust or distrust is determinant in the relationship between the two agents in the future. If, in the previous example, the 10% growth in sales was the outcome of a negotiation process where the General Manager and the Marketing manager had a frank discussion and agreed about the objective, and have a frank discussion and come to agree later on in their evaluation of what actually happened, this will increase mutual trust between them. If, on the contrary, the 10% objective was imposed by the General Manager without even hearing the Marketing Manager's opinion, and then they fight each other about who is to blame for what happened, this will decrease mutual trust and make things difficult for the future.

In summary, an 'Active Agent', in order to obtain some desired results through a decision that has to be put into practice through someone else, called a 'Reactive Agent', needs to consider three criteria: in order to establish a relationship between them that does not decrease in effectiveness through time (and can therefore be long-lasting):

- 1) the *effectiveness of the decision*, i.e., the degree to which the decision obtains the desired results
- 2) the *Reactive Agent's operational and evaluative learning*, which determines the effectiveness of future interactions in an instrumental way, since the Reactive Agent is the means by which the Active Agent obtains specific results. Particularly important is the Reactive Agent's evaluative learning with respect to trust and cooperation with the Active Agent.
- 3) *The Active Agent's own operational and evaluative learning*, which determines the action plans for the future and the initiatives of action to be taken by that agent. Again, the evaluative learning of the Active Agent with respect to trust and cooperation with the Reactive Agent is particularly important.

This three-criteria decision-making process obviously adds some complexity to the usual analysis based only on effectiveness, and is therefore not without a cost. Besides the cost of complexity, putting it into practice may often mean making sacrifices in the short run, both on the part of AA and RA, expecting a better relationship in the future that enhances effectiveness in the long run. But these criteria may be taken into account for other motives besides enhancing long-run effectiveness. These other motives will be analyzed in the following section.

Types of Motives

The behavioral literature has long since distinguished between ‘extrinsic’ and ‘intrinsic’ motives (see, e.g., McGregor, 1966, pp. 259 ff.). ‘Extrinsic motives’ are those that push the individual towards pursuing the external, explicit, tangible results that typically trigger the action. Achieving these results, and, thus, satisfying those motives is what we have called ‘effectiveness’.

‘Intrinsic’ motives are the ones that have to do with the work itself, one’s own learning, development and self-actualization. People may be interested in this partly because of improved effectiveness in the long run, but also because they think their personal development is important in its own right. Thus, they are closely related to AA’s learning.

In many formulations, intrinsic motives include motives that are pro-social or altruistic (e.g., Frey and Meier, 2002, Frey, 2003). Here we are going to group them apart, because of two reasons. First, such motives are related to RA’s learning and thus fit in very well with the model presented here. But second, because they constitute the kind of motives that will allow us to incorporate the ethical dimension in the decision-making process. In the Pérez López formulation they were called ‘transcendent’ motives, because they refer to values that transcend the individual person.⁵ These motives indicate a genuine interest in the development and motives of the other person (his/her learning) that goes beyond considering exclusively future effectiveness.

Having only motives of the extrinsic type results in neglecting first the learning of RA, who would then be less able or less willing to cooperate in the future; which would make the same problem (or a similar one) more difficult to solve next time. It also results in neglecting one’s own learning, since obtaining immediate results may make future results worse than they could be with an appropriate learning process; and this will make it more difficult (relatively speaking) to solve similar future problems. The only way for a decision-maker to solve a problem and not make it more difficult in the future is to care about intrinsic and transcendent motives.

Notice that ‘transcendent’ motives go well beyond the usual meaning of ‘altruistic’, ‘philanthropic’, ‘contributive’, or ‘benevolent’. First, they incorporate rational motives and not only the ‘taste’ for emotional benevolence. AA may not ‘feel like’ doing something for RA, but may consider that that’s exactly what he/she should rationally do to have in the future a continuing, mutually beneficial relationship. Second, at a different level, AA may take into account the effects on the other agent not because of an ongoing relationship, but because of a moral obligation. Sen (1977), when analyzing the actions taken by an individual in favor of another one, distinguished between “sympathy” and “commitment”. Sympathy exists when one person is affected in his/her own feelings when something happens to another (and, thus, in economic language, the welfare of other people is included in one’s own utility function); while commitment occurs when what happens to someone else does not affect a given person directly, but that person thinks it is his/her moral obligation to do something about it. His example is that of a person seeing a child being tortured: if this makes this person sick, that’s ‘sympathy’; if, in contrast, this does not affect that person in his/her feelings, but the person

⁵ Interestingly, it is the same word used by William George to indicate a type of leadership with very similar values (George, 1999, 2001, 2003).

thinks it is his moral obligation to try to stop it, that is 'commitment'. 'Transcendent' motives are closely related with Sen's 'commitment', and, thus, to ethical behavior. The concept of Frey and Meier (2002) of 'pro-social' intrinsic preferences is also similar to that.

Long-run effectiveness is not guaranteed unless intrinsic and transcendent motives exist and are independent of the desire for future effectiveness, because the development of trust between two people is based on their mutual appreciation of sincerity and genuine interest. Therefore, if we want to guarantee long-run effectiveness, we need motives other than merely extrinsic ones.

Rationality and Virtuousness

The learning of the two agents may not take place immediately or be seen right away. Some 'tricks' may induce a given RA into cooperation in the short-run, but as soon as they are discovered and RA's *learns*, the tricks cease to be effective. Similarly, AA may find it difficult to initiate some course of action because of the effort required, but find at the end (it might even be a long time down the line!) that it was worth it. Therefore, the approach introduced here should not be considered as something intended to satisfy only immediate or perceived needs, but to also satisfy the (boundedly rational) long-term needs of individuals considered as human beings. Indeed, short-term effectiveness is not the main objective of this approach at all. As we will attempt to show in this section, it rather intends to promote *rationality* and *virtuousness*. Let's analyze these two concepts next.

'Irrational' behavior is behavior that does not take into account all the likely consequences of one's actions, and is guided only by perceived needs, short-term impulses, emotions, or visceral factors (Lowenstein, 1996), disregarding other long-term aspects that may become more important in the end, like self-actualization, achieving one's potential, achievement, or simply personal growth. 'Rational behavior', in contrast, takes into account all the effects of one's action, including short-run and long-run effects, even if one has to sacrifice some immediate advantages or do something that might be unpleasant. Irrational behavior is typically based in direct experience only, while *rationality* is based on the abstract knowledge that human beings alone are able to possess. The analysis above, thus, has to be performed in terms of rationality, including the consideration of the long-run consequences and not only of the immediate ones, and by using abstract knowledge. Analyzing only the immediate consequences would be equivalent to forgetting about the possible learning of the two individuals involved, since in the very short run, this learning does not matter at all.

Rationality, however, is not enough to make decisions that are consistent with the model presented. Besides rationality, agents need to develop *virtuousness*, i.e, the willpower that allows them to do what is rational in spite of the short-run sacrifices mentioned. And, according to Aristotle, virtues are acquired only by practice. Thus, the application of our model requires another crucial aspect of ethics. This, as we will see next, has important implications for organizational decision-making.

Organizational Decision-Making

The above analysis applies, rigorously speaking, only to the interactions between two human beings. However, we can extend it (by analogy at least) to organizations or groups of people, including all stakeholders involved. Thus, AA can be 'the firm', and RA its 'employees', or its

'customers', or even abstract concepts such as 'the market'. AA can also be the CEO (or any manager at any level) and RA can be the organizational unit below that person (i.e., 'the firm', 'the business unit', the 'department', and so on). Actually, any collective entity like 'the firm', 'the market' is represented in any given situation by a specific person, like a salesperson may represent 'the firm', a specific customer may represent 'the market', or a department's manager may represent 'the department'. In any case, our approach emphasizes the importance in the analysis of decisions of learning and of the different aspects of that learning (individual and collective) at all levels.

The familiar analysis of decision-making suggested in the marketing, finance, production or cost accounting textbooks is typically limited to the effectiveness of decisions, i.e., the explicit results obtained and sought for: the impact of the decision alternatives under consideration on short-run profit, or even on a more immediate variable like contribution. At most, if the decision clearly affects several periods in a way that can reasonably be estimated, a present value analysis is (correctly, although that analysis is not without some limitations) supposed to take into account the impact of the decision in the value of the firm. But beyond that, in most cases, while it is obvious that the learning of the organization will have a strong influence on the long-run cash flows, it is not possible to evaluate that influence with any degree of precision. Learning might thus be overlooked or underestimated.

In our analysis, we have shown the need to consider the learning of both parties in any interaction, if we want the relationship between those two parties to continue indefinitely, and hopefully improve. Actually, leaving learning aside is equivalent to thinking of people as merely mechanical devices. That is exactly what happens in most economic models, where people are represented by utility functions that have well-defined preferences between any two goods, very precise (subjective) estimates with respect to uncertainty, and precise knowledge about cost and revenue functions. If this representation were realistic, it would be difficult to argue against firm value as the overriding goal of the business firms, as economic models show quite clearly. Jensen possibly expresses the argument in the simplest terms. Starting from a scenario where all production runs are infinite and cash flow streams level and perpetual, he briefly shows how profit maximization is socially optimal:

"In this simple situation, a firm taking inputs out of the economy and putting its outputs of goods and services back into the economy increases aggregate welfare if the prices at which it sells the goods more than cover the costs it incurs in purchasing the inputs. Clearly the firm should expand its output as long as an additional dollar of resources taken out of the economy is valued by the consumers of the incremental product at more than one dollar. Note that the difference between these revenues and costs is profits. This is the reason (under the assumption that there are no externalities) that profit maximization leads to an efficient social outcome." (Jensen, 2000, p. 43)

Of course, the initial assumption of level and perpetual cash-flows is rather innocuous: if this does not happen, present value analysis leads to the same conclusion. But once we consider the lack of knowledge that people have to evaluate the impact of current decisions in the firm's future, learning becomes a major objective. Senge states it quite clearly:

"This is why, for me, learning takes precedence over optimizing. The fundamental challenges of management are not about 'finding the right answer' so much as about seeking better understanding of the consequences of our actions, and especially of how we may be unwittingly heading in directions opposite to what we intend. (...) Thus, whether intended or

not, firm value maximization will almost always become, by default, short-term profit maximization. The reason is that the complex feedback dynamics that bedevil our simplistic causal theories take time to play out. Given a short enough time horizon, many of these feedbacks can be ignored. This is why manipulating profits over the short term is much easier than building wealth over the long term. A simple causal theory may suffice for the former, while the latter requires a much more complex causal theory.” (Senge, 2000, p. 63-65)⁶

In spite of this, in some situations, analyzing decisions by focusing on the short run, quantitative variables and their impact on the firm’s value may just be the right thing to do. Some (or perhaps even ‘many’, in quantitative terms) decisions affect only the short-run and have very little impact on learning: the contribution analysis in any cost accounting textbook typically analyzes situations of this kind. Also, as stated above, when the decisions do have an impact on the long run, but this impact can be reasonably estimated, the familiar textbook present value calculations may be all is needed. But in most ‘high-level’, strategic or organizational decisions, characteristic of the management professions, learning is central to the problem. Thus, the importance of learning is often greater than the quantitative analysis.

Unfortunately, as stated above, very often the learning dimensions are simply forgotten or overlooked. The emphasis in recent times on short run and immediate results at the expense of long run, intangible variables, has contributed to this state of affairs. While in theory long-run effects can be thought to be included in present value analysis, in practice the impact of organizational learning on long-run profits, cash-flows or value of the firm is completely impossible to evaluate, even by a rough approximation; which might just be a different way of restating the Senge quote above. As we will show next, our previous framework adds to his analysis a structured way to include learning in the analysis of specific decisions. Let us see how.

Two Stylized Examples

I will next illustrate this by expanding on two examples mentioned in the preceding section, i.e., that of ‘the firm’ as AA and ‘the customers’ as RA, and that of ‘a manager’ as AA and ‘the organizational unit below’ as RA. The first one represents the problems that have to do with the relationship between the firm and the market (i.e., market strategy); while the second has to do with the decisions that are internal to the organization (i.e., institutional strategy). These two examples, representative of the two main endeavors of the firm (external and internal), we be taken as building blocks of a theory of management and organization.

‘Firm’ as AA and ‘customers’ as RA

Let’s first assume that AA is ‘the firm’, and RA is ‘the set of customers’. Interactions between those two collectives may have to do (for instance) with decisions on advertising, pricing, sales efforts, distribution channels and so on.

⁶ A word of caution is needed at this point. Our model goes well beyond Senge’s argument, because he seems to refer only to the problem of mathematical complexity by asking for a ‘much more complex causal theory’. In contrast, in the approach followed in this paper, a causal explanation, no matter how complex, is not enough to explain human phenomena and the different kinds of learning that we have explored. Rather, an intentional explanation is needed (Elster, 1983; Ghoshal, 2005). From this point of view, Senge seems to be analyzing a simpler problem; but this can make Senge’s case in asking for giving more importance to learning only stronger.

When the above analysis on the interaction between two people is applied to that situation, any decision made by 'the firm' interacting with 'customers' has to take into account:

- a) The effectiveness of the decision (typically, in economic terms)
- b) The learning of the customers
- c) The learning of the firm

We study them in turn.

First, effectiveness is an obvious criterion: as stated above, it is what the typical elementary textbook analyzes, and it first has to do with the basic economic variables like 'revenues', and then perhaps their expected consequences in terms of 'profit', 'shareholder value', 'value of the firm' or any alternative formulation. The 'effectiveness' of every specific decision is the degree to which the goals of the firm in terms of such variables are attained.

Second, customers' learning in a purchasing interaction is essentially evaluative: their (conceivable) operational learning reduces to the search for the product and/or the physical operation of actually purchasing it.

From an evaluative point of view, then, consumers will learn in two dimensions:

- a) whether the product satisfies their real needs, both in the short and in the long run, and whether it has undesirable, perhaps unforeseen consequences that they will try to avoid from then on.
- b) whether the firm can be trusted to help them satisfy their needs, i.e., whether the firm has been truthful about the product, whether it has advised them adequately about their possible choices of specific products, whether it has provided an adequate service, whether it has helped to solve possible problems after the sale (post-sale service), and so on.

The first of these two dimensions partly depends on the second: the customer may be happier with the product if the firm has been helpful and cooperative; and, this is only possible if the firm is able to make some short-term sacrifices in order to increase customer satisfaction. If all the firm wants is getting customer orders in the short run, and forgets about all these other aspects, real customer satisfaction (not always reflected in the usual questionnaires used to measure it) cannot be high. And this, of course, will have clear implications in the long-run.

Finally, the firm's learning will be both operational and evaluative. First, by doing something in order to get an order delivered, from procurement to production and from production to marketing, the firm will learn operationally how to do the same operations better or more efficiently.

But second, and more important, the firm will learn to satisfy the needs of the customer better. When customers order a product, they are trying to satisfy some underlying needs that are typically complex, having more than one dimension. For instance, if they buy bread they are trying to satisfy an (obvious) physiological need for food (calories); but at the same time they want some taste that they can enjoy, a type of food that they can combine with other types (e.g., in a sandwich), some food adequate for a social occasion, and so on. Some needs may be immediate and perceived (weight, size and its relationship with how hungry they are); others are just as real but unknown at the time

of the purchase (taste, health properties, etc.); others may be obvious to customers, but unknown to the firm. The firm can then make a product that is intended to satisfy the real (complex) needs of the customers, or only their obvious, perceived needs at the time of the purchase. To the extent that the firm is in contact with customers, it can learn what their real needs are, and attempt to satisfy them by developing a *distinctive competence* (Selznick, 1956) as opposed to simply obtaining orders based on the obvious, perceived needs; or, even worse, based on 'tricks' of different kinds. But it is important to stress that the firm's learning about what the real needs of the customers are and what products that are feasible can satisfy them can take place only if the firm's intention is to do precisely that. A strategy of opportunistic adaptation (giving customers what they explicitly ask for) cannot produce such a result: the firm will always 'lag behind' real customer demand and the competitor's actual supply, and will only see the obvious. The firm builds a distinctive competence only when it learns to satisfy some types of consumers needs; and, as Selznick already pointed out in his 1956 classic, a strategy of opportunistic adaptation destroys this distinctive competence.

In summary, the firm's learning has two different aspects: the operational development of knowledge and skills of individuals, and the development of their willingness and capacity to satisfy customer needs. As we shall show next, this is closely related to the organization problem, i.e., the network of relationships between the people that actually work for the organization; but in any case, this constitutes what can be considered the 'external mission' of the firm⁷, in the Pérez López terminology: to satisfy some type of customer needs. By considering this as a decision criterion, the firm learns operationally to become an expert in satisfying such needs, and, thus, acquires a competitive advantage by developing good relationships of mutual trust between itself and its customers.

'A manager' as AA and 'an organizational unit' as RA

In our second example, AA is a manager, and RA the organizational unit below that manager. The decisions that can be represented by this situation are of course many: any time an employee receives an order from his/her boss we have a situation of this kind. Giving an employee the assignment to perform a specific task, or to implement a decision made by the manager are typical examples.

Again, we will consider the three criteria for decision-making:

- a) The effectiveness of the decision
- b) The manager's learning
- c) The unit's learning

In this case, the effectiveness of the decision may not have to do with immediate economic variables, but with tangible, and possibly measurable aspects of performance. In any case, effectiveness in this context means achieving specific results pursued with the decision, and is obviously an important criterion.

⁷ Incidentally, this is again parallel to the William George strategy (George, 1999, 2001, 2003).

The manager's learning will be partly operational and partly evaluative. From an operational point of view, a manager can learn about the best way to solve operational problems and use the appropriate techniques: 'what works' and what 'does not work', and how to handle the relationships with the employees. From an evaluative point of view, the learning is all the more important, because the manager learns to command his/her people and develops his/her relationships with them, to what extent he/she can trust them, creating a 'good' or 'bad' climate in the unit. The importance of this for any organization cannot be underestimated.

The unit's learning will also be partly operational and partly evaluative. In terms of the operational part, the above description also applies. The evaluative part is more interesting because it involves the relationship between the employees and their manager and their understanding of their job. The role of the manager will essentially consist of making them understand the importance of what they are doing, and the importance of doing it right. If properly managed, employees will understand this and will attempt to do it in a way that ultimately satisfies the real needs of customers. This will mean a sense of unity in the organization and will allow each employee to identify with the organization's external mission. This is what Pérez López (1993) called "internal mission".

One specific aspect of learning that is particularly important in this context is individuals' learning about each other and building or destroying trust within the organization. A necessary condition for the organization's long-run survival is employees' willingness to cooperate with each other to solve the firm's problems. We argued in our analysis of the symmetrical agency problem that in a two-person relationship the development of trust is crucial for the future of the relationship; quite obviously, it is therefore all the more important in the context of an organization. The future of the organization depends on building trust in such a way that the individual can trust the organization (which, of course, is a complex interrelationship between many individuals) and the organization can trust the individual. In this way, individuals can internalize the goals of the organization and identify with and be loyal to it, as Simon's classic analysis showed (Simon, 1947, Ch. 10; Rosanas and Velilla, 2003). The cornerstone of that unity is the relationship between the manager and the people in the organizational unit below him/her.

In Summary: Three Criteria for Decision-making

The previous analysis leads naturally to a way of analyzing decisions in organizations that does not limit itself to the textbook analysis of effectiveness. We have somewhat artificially separated, 'internal' and 'external' decisions; in practice, almost all decisions in organizations have to do at the same time with the firm relationship with the market, and with the internal organization of the firm. Hence, it becomes important to put the criteria of the two previous examples into an integrated framework. In fact, and given the logic of the whole analysis, coming from our 'symmetrical' and 'dynamic' agency theory, this is not a difficult task. It reduces, for any decision-making process within the context of an organization, to considering three criteria again:

- 1) effectiveness: the impact on immediate, explicit, measurable variables, of which the economic ones are particularly important. Contribution, profits, shareholder value, firm's value or whatever are the most common expressions, depending on the problem at hand. Quite obviously, it would be absurd for a business firm to forget about such variables. This is related to the satisfaction of the extrinsic motives of producers and consumers as well, and is the only aspect considered in reductionist models of the firm based strictly on economics as a discipline and as a theory.
- 2) the improvement of the firm's distinctive competence, or the contribution of the chosen course of action to the development of its external mission. This goes beyond the immediate economic variables by defining the type of real needs of customers that the firm intends to satisfy. It is related to the intrinsic motives of the producers, their operational and evaluative learning, and the development of their capacities to solve the customers' problems.
- 3) the unity of the organization: identification with the organizational objectives and with the other employees is the third criterion to take into account. The degree to which producers are identified with the organization is closely related, of course, to the degree to which they are identified with the organizational goal of satisfying customers' needs.

As we have attempted to show, the three criteria are interrelated. Long-run effectiveness depends on distinctive competence and on unity, distinctive competence depends on unity, and unity depends on distinctive competence. But a reductionist approach that takes into account only one of the criteria (perhaps long-run effectiveness, measured by shareholder value, profit, or any other financial variable), and considers all the others subsumed in that one won't do, and will make the Senge quotation above and the comments that follow right. Distinctive competence and unity are variables that have to be pursued on their own merit, not only as instrumental to effectiveness, because otherwise the uncertainties are such that the short-run financial variables will prevail.

In summary, the approach described in this paper goes beyond the usual reductionism of economic variables, and attempts to initiate a different one with a humanistic basis, that explicitly considers the development of human beings in organizations and the social role organizations fulfill in society. Thus, it can be considered a possible beginning to develop the humanistic approach Melé called for (2003). Looking at the future, this approach will need to make progress both in the directions of further structuring and formalizing the theory, and integrating it further with other disciplines, among which ethics is, of course, possibly the most important.

References

- Anderson, C.: 1997, 'Values-based Management', *Academy of Management Executive*, **11** (4), 25 - 46.
- Argandoña, A.: 2007, 'Integrating ethics into action theory and organizational theory' *Journal of Business Ethics* (this issue).
- Davis, J. H., F. D. Schoorman, and L. Donaldson: 1997, 'Toward a stewardship theory of management', *Academy of Management Review*, **22**, 20-47.
- DiMaggio, P. J., and W. W. Powell: 1991, 'Introduction', In W.W. Powell and P. J. DiMaggio (eds.), *The New Institutionalism in Organizational Analysis* (University of Chicago Press, Chicago).
- Donaldson, L.: 2002, 'Damned by Our Own Theories: Contradictions Between Theories and Management Education', *Academy of Management Learning and Education*, **1** (1), 96-106.
- Elster, J.: 1983. *Explaining Technical Change*, Cambridge, England, Cambridge University Press.
- Elster, J.: 1994, 'Functional Explanation in Social Science', in Martin, M., and McIntyre, L. (ed.), *Readings in the Philosophy of Social Sciences* (The MIT Press, Cambridge, Massachusetts).
- Ferraro, F., J. Pfeffer and R. Sutton: 2005, 'Economics Language and Assumptions: How Theories May Become self-Fulfilling', *Academy of Management Review*, **30** (1), 8-24.
- Frederick, R. E.: 2002, 'An outline of ethical relativism and ethical absolutism', in R. E. Frederick (ed.), *A Companion to Business Ethics* (Blackwell, Oxford), pp. 65-80.
- Frey, B. and S. Meier: 2002, 'Pro-Social Behavior, Reciprocity or Both?', CESIFO, Munich, Working Paper No. 750.
- Frey, B.: 2003, 'Corporate Governance: What Can We Learn from Public Governance?' Working Paper, Institute for Empirical Economic Research, University of Zurich.
- George, W. W.: 1999, 'Mission Driven, Values Centered', *Executive Excellence*, **16** (8), 6, August
- George, W. W.: 2001, 'Medtronic's Chairman William George on how mission-driven companies create long-term shareholder value', *Academy of Management Executive*, **15** (4), 39-47.
- George, W. W.: 2003, *Authentic leadership: rediscovering the secrets to creating lasting value* Jossey-Bass, San Francisco).
- Ghoshal, S.: 2005 'Bad Management Theories are Destroying Good Management Practices', *Academy of Management Learning and Education*, **4** (1), 75-91.
- Grossman, S. and O. Hart: 1983, 'An Analysis of the Principal-Agent Problem', *Econometrica*, **51** (1), 7-45.
- Holmström, B.: 1979, 'Moral Hazard and Observability', *The Bell Journal of Economics*, **10** (1), 74-91.
- Jensen, M.: 2000, 'Value Maximization, Stakeholder Theory and the Corporate Objective Function', in M. Beer and N. Nohria, (eds.), *Breaking the Code of Change* (Harvard Business School Press, Boston). Also in *Journal of Corporate Finance*, Fall 2001.
- Jensen, M.: 1983, 'Organization Theory and Methodology', *The Accounting Review*, LVIII (2), pp. 319-339

- Jensen, M. and B. Meckling: 1976, 'Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure', *Journal of Financial Economics*, **3**, 305-360.
- Lowenstein, G.: 1996, 'Out of Control: Visceral Influences on Behavior', *Organizational Behavior and Human Decision Processes*, 1996. **65** (2), March, 272-92.
- March, J. and J. Olsen: 1975, 'The Uncertainty of the Past: Organizational Learning under Ambiguity', *European Journal of Political Research*, **3**, 147-171.
- Marschak, J, and R. Radner: 1972. *Economic Theory of Teams* (The Cowless Foundation, New Haven and London)
- Melé, D.: 2003, 'The Challenge of Humanistic Management', *Journal of Business Ethics*, **44** (1), 77-88.
- Meyer, J, and B. Rowan: 1977, 'Institutionalized Organizations: Formal Structure as Myth and Ceremony', *American Journal of Sociology*, **83**, 340-363
- McGregor, D.: 1966, *Leadership and Motivation* (MIT Press, Cambridge, Mass.).
- Nilakant, V. and H. Rao: 1994, 'Agency theory and uncertainty in organizations: An evaluation', *Organization Studies*, **15** (5), pp. 649-72.
- O'Donoghue, T., and M. Rabin: 1999, 'Doing it Now or Later', *American Economic Review*, **89** (1), pp. 103-124, March.
- Pérez López, J. A.: 1991, *Teoría de la Acción humana en las organizaciones* (Rialp, Madrid, Ediciones).
- Pérez López, J. A.: 1993, *Fundamentos de la Dirección de Empresas* (Ediciones Rialp, Madrid).
- Rosanas, J. M.: 2004. *Bounded rationality, value systems and time inconsistency of preferences as rational foundations for the concept of trust*, IESE Business School, Working Paper no. 567.
- Rosanas, J. M., and M. Velilla: 2003, 'Loyalty and Trust as the Ethical Bases of Organizations', *Journal of Business Ethics*, **44**, 49-59.
- Ross, S.: 1973, 'The Economic Theory of Agency: The Principal's Problem', *American Economic Review*, **63** (2), 134-39.
- Scott, R.: 1987, *Organizations: rational, natural and open systems*, 2nd (ed.), (Prentice-Hall, New York)
- Selznick, P.: 1957, *Leadership in Administration* (University of California Press, Berkeley Los Angeles and London).
- Selznick, P.: 1996, Institutionalism, 'Old' and 'New', *Administrative Science Quarterly*, **41**, 270-277.
- Sen, A.: 1977, 'Rational Fools: A Critique of the Behavioral Foundations of Economic Theory', *Philosophy and Public Affairs*, **6**, pp. 317-44.
- Senge, P.: 2000, 'The Puzzles and Paradoxes of How Living Companies Create Wealth', in M. Beer and N. Nohria, eds., *Breaking the Code of Change* (Harvard Business School Press, Boston) pp. 59-81.
- Simon, H.: 1947 first edition, 1997 fourth edition. *Administrative Behavior* (The Free Press, New York)
- Yariv, L.: 2002, 'I'll See It When I Believe It - A Simple Model of Cognitive Consistency', *Cowles Foundation Discussion Paper*, No. 1352.