

# Human Nature and Institutional Analysis

Benito Arruñada<sup>1</sup>

Universitat Pompeu Fabra. Barcelona

Published in E. Brousseau and J.-M. Glachant, eds., *New Institutional Economics: A Guidebook*, Cambridge University Press, Cambridge, 2008, pp. 81-99.

## I. Abstract

This chapter reviews some findings in cognition sciences and examines their consequences for the analysis of institutions. It starts by exploring how humans' specialization in producing knowledge ensures our success in dominating the environment but also changes fast our environment. So fast that it did not give time to natural selection to adapt our biology, causing it to be potentially maladapted in important dimensions. A main function of institutions is therefore to fill the gap between the demands of our relatively new environment and our biology, still adapted to our ancestral environment as hunter-gatherers. Moreover, institutions are built with the available elements, which include our instincts. A deeper understanding of both aspects, their adaptive function and this recruitment of ancestral instincts, will add greatly to our ability to manage institutions.

JEL Classification: D01, D02

Keywords: evolution, biology, behavior, institutions.

---

<sup>1</sup> Department of Economics and Business, Universitat Pompeu Fabra. Mail: Trias Fargas, 25. 08005-Barcelona, Spain. E-mail: [benito.arrunada@upf.edu](mailto:benito.arrunada@upf.edu). The author thanks Jean E. Ensminger, Jean-Michel Glachant, Geoffrey M. Hodgson, Paul H. Rubin, Marta Serra, Robert L. Trivers, Xosé H. Vázquez, David Sloan Wilson and participants at the European School of New Institutional Economics, The Ronald Coase Institute, Universidad Pablo de Olavide, Université de Paris I, Université de Paris X, Universidades de Castilla y León and the Gruter Institute for Law and Behavioral Research, for their comments; and, for financial support, the MCYT, an agency of the Spanish Government through grant SEJ2005-03871/ECON and the European Commission through the Integrated Project CIT3-513420. Usual disclaimers apply.

## **II. Introduction**

The human mind was mainly designed in a competitive process of natural genetic selection, which is characterized by random genetic mutation, producing new traits, and cumulative selection of those traits that allow individuals who carry them to survive and reproduce more. Natural selection thus acts as a chief design engineer, even if other forces, such as sexual selection, path dependency and simple noise, are also present. We see well now only because a long series of mutations triggered redesigns which permitted our ancestors' sight to improve. The same happens with our mental processes, even those considered more rational, involved in making decisions and interacting socially.

Modern cognition sciences perform a sort of “reverse engineering” of these mental processes. Their findings may trigger a scientific revolution of Copernican proportions in the social sciences and in any case require a full reconsideration of standard assumptions about human behavior, related to both rationality and cooperation.

This chapter reviews some of these findings and examines some of their consequences for the analysis of institutions and organizations. We start by exploring the consequences of our specialization in producing knowledge, which are twofold: it has ensured our success in dominating the environment but has also changed the environment very fast and radically. This change occurred so fast that it did not allow time for natural selection to adapt our biology, causing us to be maladapted in important dimensions. To adapt, we therefore need the artifacts we call institutions. A new view of institutions thus emerges, which sees their function as that of filling the gap between our biology, which is still adapted to our ancestral environment as hunter-gatherers, and the demands of our relatively new environment.

The development of institutions therefore facilitates cooperative transactions which seem to rely less on our instinctual psychology and more on artificially designed structures of enforcement. These artificial systems rely on instincts, however, as they recruit them for performing new functions within the institutional arrangement.

Consequently, cooperation is not only grounded on a calculation of costs and benefits, as it is sometimes assumed in utility maximizing models of human behavior. Automatic mechanisms, evolved in ancestral environments, play essential roles, and their functioning has to be understood for wisely structuring our artificial enforcement systems (including those of firms), as well as for using our calculative rationality successfully when we interact with other individuals.

This emphasis on instincts ties in with classic and institutional economics. For instance, Adam Smith saw humans as essentially instinctive (Coase, 1976) and he correctly understood that instincts (that is, his “sentiments”) are adaptive and, under normal circumstances, make no mistake. This is why affairs of survival importance, such as self-preservation and reproduction, are not “entrusted to the slow and uncertain determinations of our reason” but to “original and immediate instincts” (Smith, 1759, II, iii). Instinctive Darwinian psychology was also important for institutional analysts, such as Veblen (Hodgson, 2004a and 2004b), because they saw habits as the basis of human behavior, and habits are close to current views of the mind, based on modules and heuristics.

This chapter will proceed in four stages. First, it will examine how the specialization of human beings in cognition leads to a modular design of the human mind and how it grants both biological success and maladaptation. Next, it will explore the consequences of this view, in terms of modular instincts and environmental maladaptation, for the two key behavioral assumptions, those of rationality and cooperation. Then it will explain how institutions allow us to fill the gap in our innate maladaptation, a job for which institutions often recruit instincts originally designed for other purposes. A final section concludes.

### **III. Consequences of our cognitive specialization**

#### **III.A. Human beings are specialists in the *cognitive niche***

We are not very good at flying but we do build planes that fly faster than any animal. Already in ancestral times, we were the best predators: thanks to our hunting technologies, both physical

and social, we were able to hunt animals that were too big to be hunted by any other predator. We achieved it all by becoming knowledge specialists, by entering the “cognitive niche” (Tooby and De Vore, 1987), and developing increasingly sophisticated tools, with which we have radically changed our environment.

This specialization in cognition and technology constrains our design but also explains our dominant position in nature. Our design is constrained to have certain physiological and social constraints that make brain development possible, and, most importantly, to have a modular mind instead of a general processing mind. Cognitive specialization also brought about a substantial comparative advantage over our animal prey and competitors, with the side consequence that we also became maladapted to the rapid changes we cause in our own environment.

### **III.B. Modular design**

Firstly, an intelligent mind has to be produced and has to function economically, because brains are very costly to operate: our brain weighs only 2% of our total body weight but it spends around 20% of our energy. To be efficient, cognitive specialization requires a certain degree of modularity in the internal workings of our mind, as any other complex system (Simon, 1962). Otherwise, a general-purpose mind would have to use the same methods and tools for different problems which present different information structures. In contrast, specialized modules make it possible to optimize the use of the information available in the environment. With this specialization, the mind contains mechanisms that are, in a sense, “better than rational” because they minimize the use of information, speed up decisions and produce sophisticated solutions (Cosmides and Tooby, 1994). We will see below that instincts provide speedy optimal solutions without conscious rational thinking and emotions even achieve optimal strategic self-commitment by suspending rationality.

### **III.C. Success—and maladaptation**

Second, intelligence confers on human beings a huge comparative advantage over most other animals, because we are able to develop new technologies, including weapons and hunting

techniques, faster than our prey and competitors can evolve defenses against them, given that they develop them only by natural selection. We therefore become dominant and much of our prey tends to go extinct, as shown by the now abundant record that we exterminated big animals each time we first arrived on a continent.

However, our cognitive specialization has the paradoxical consequence that we outrun not only our prey but also ourselves. In the last ten thousand years (an instant for natural selection), we have also changed our own environment far faster than our own genetics could adapt to such change. Natural selection is powerful but slow, requiring thousands of generations. The human brain thus evolved under the selective pressures faced by our ancestors in the ancestral “environment of evolutionary adaptedness” of the Pleistocene period (1.8 million to ten thousand years ago), the only interval long enough to allow significant genetic adaptation.

Our mind is therefore designed to cope with the problems relevant for survival and reproduction at that remote time—those of habitat selection, foraging, social exchange, competition from others, contagion avoidance, and sexual rivalry. Our mental hardware is fine-tuned to live in small nomadic tribes, hunting and gathering fruits, in a world with few technologies—just fire as well as stone and wood tools, and little interaction and trade outside the group.

In a manner consistent with our cognitive specialization, the solution for this maladaptation has also been technological: we use institutional “technologies” to improve our fitness. The function of institutions is therefore to enhance our capacity to reason and interact, allowing us to overcome our own evolutionary constraints, mainly through self and social control.

## **IV. Rationality**

When engineers started to design mechanical robots, they soon realized that seemingly simple tasks, such as recognizing objects, are instead tremendously complex, and achievements in such tasks have been slow. Computers are now very good at using mathematical and logical rules, and one of them even beats the world chess champion. At the same time, experimental psychologists have shown that human beings err systematically in simple logical problems and

poorly assess the probability of individual events. Why is our mind so powerful and at the same time so limited? Why do bumblebees perform better than most humans at probabilistic induction? The answer is simple. Our mind is *powerful but economical* in the use of resources. It spends resources in solving problems that were relevant for survival in our evolutionary past, but it does not care about those which were irrelevant.

Make no mistake. The human mind is very *powerful* indeed. It routinely, even effortlessly, manages to solve the most difficult problems: those without solution, such as identifying the factors in a product—so-called “ill-posed” problems. It is so powerful that we are *better than rational* on evolutionary recurrent tasks, such as recognizing objects, acquiring grammar, or comprehending speech. Robot designers soon realized how difficult it was to endow their creatures with the crudest rudiments of sight or walking. After several decades of research, even purposely-designed robots are only able to walk clumsily, or to identify only the set of forms that they have been programmed to “see.”

But our mind is also *economical*, meaning that it uses only those resources required to succeed in a given environment. Our mind does not produce “scientific” solutions, with general validity, but solutions which are good enough to master the local environment. Our rationality is bounded not only because it is subject to constraints, but because it is developed and adapted to certain environments: it is *ecological*, meaning that it is adapted, first, to our common ancestral environment of evolutionary adaptation and, second, to our learning environment, probably with much more malleable consequences.<sup>2</sup>

Many observed decisional “failures” in the artificial environments of experimental psychology and economics may therefore be a consequence of its artificiality, its absence in the natural environment. For example, humans “fail” when applying probability theory. For instance, we fall prey to the “gambler’s fallacy,” feeling that, for example, lottery numbers with all their digits repeated (e.g., 33333) have a lower probability of winning than numbers with variable

---

<sup>2</sup> The idea that individuals (and organizations) decide by using heuristics which work relatively well in a given environment, be it natural or social, was proposed by Simon (1956). Reference to “ecological rationality” is found in Tooby and Cosmides (1992). See also Gigerenzer and Todd (1999), who stress how the mind makes efficient use of the structure of information available in the environment and Smith (2003) for a view from experimental economics.

digits (53487). But it might well be the case that our mind is adapted to environments in which very few events are independent and most variables are correlated. In nature, very few successes are independent and it is unclear how many there are in the current world, perhaps with the exception of some casinos and stockmarkets. Independence is often open to question. After an aircraft crash, most people are more afraid to fly. That would not make sense if aircraft crashes were independent events, but are they really independent?

Similarly, our mind seems to have developed an ability to process probabilities in terms of relative frequencies in the long term, not as numbers expressing confidence in a single event. This explains that, when probabilities are presented as frequencies (“one out of one hundred is sick”) instead of single events (“probability that John is sick”), people are much more accurate. This “frequentist” view of the mind somehow dilutes the claims about several alleged biases and fallacies, such as the overconfidence bias, conjunction fallacy, and base-rate neglect.<sup>3</sup>

#### **IV.A. Instinctive rationality**

During the last few centuries, we have become accustomed to separate reason and emotions, also considering emotions as inferior to reason. This Cartesian separation of emotions and reason is seriously flawed, however. Our mind relies on instinctive mechanisms, including emotions, to solve the most relevant problems, those on whose correct solution hinged our survival and reproduction. Furthermore, emotions are a necessary ingredient of rationality: ill people who have lost part of their frontal lobe are “perfectly rational” but their loss of emotions seems to damage their decision-making capacities. They are often incapable of deciding and, instead, keep evaluating advantages and disadvantages without ever reaching a conclusion (Damasio, 1994).

The adaptive consequences of emotions are obvious in simple ones, such as hunger, which moves us to search for food, and the pleasure of eating, which leads us to accumulate reserves in our bodies. Furthermore, emotions are often adaptive even when they seem to harm the individual, and tend, therefore, to be considered “irrational.” For example, having a hot temper

---

<sup>3</sup> See Gigerenzer (2000), whose work has been criticized, however, both in terms of its results (e.g., Kleithner *et al.*, 1997) and for distorting the position of the “biases and fallacies” paradigm (Markoczy and Goldberg, 1998, 400-402).

that leads us to react violently to even minor offenses may have a deterrent effect which helps such person in a lawless context.

More complex emotions also have adaptive value. For instance, happiness mobilizes resources to fit in our environment and to reproduce. The paradoxes of happiness may be explained from this perspective. First, we feel happy when we observe that our *relative* position is good. This seems silly but it is not, because relative positions inform us about which level of achievement we should be aiming for. Second, in determining the degree of happiness, we give greater importance to changes than to levels of achievement. This emphasis on changes renews our motivation to strive in the search for happiness and, therefore, environmental fitness—both people who win a lottery or whom suffer misfortune adapt fast to their new situation. Third, we feel more negative than positive emotions and we tend to grant greater weight to losses than to gains. This asymmetric feeling may also be adaptive because the consequences of losses and gains are intrinsically asymmetric, given that losses threaten survival while gains do not increase reproductive success in an equivalent proportion. There are, in a sense, diminishing returns to wealth in generating fitness.

#### **IV.B. Ecological rationality: the maladaptation of our instincts**

There are many examples of emotions that not longer seem well adapted, however. For instance, our feeding emotions were probably useful in our ancestral set-up, characterized by unreliable food supply, but are badly adapted for wealthy societies. Consequently, we now need to spend resources and develop self-controlling mechanisms to avoid dying earlier from overeating. Without self-control, we tend to eat too much, especially sugar, fat and salt. A taste for sweetness motivated our ancestors to eat fruits but it became maladapted when we recently developed sugar and candy, transforming the taste for sweetness into a damaging sweet tooth.

Let us examine two prominent examples of emotional maladaptation with vast economic consequences: risk aversion and weakness of will.



## **Risk aversion**

As any other essential trait of human beings, risk aversion probably has an innate component, as shown by the presence of a certain asymmetry of gains and losses and perhaps even excessive risk aversion.

Evolutionary optimal preferences about risk should be adapted to the ancestral environment in which people were living on the edge of subsistence. Under such dire straits it makes sense to evolve risk aversion preferences, especially toward losses, and this may be behind some forms of asymmetry which have been observed in experiments. Our current environment is less uncertain and its optimal rate of risk aversion may well be lower. However, natural selection is too slow to catch up with fast environmental change. Therefore, instinctive risk aversion may be leading modern humans to excessive prudence. We are risk averse to avoid risks that we are programmed to (wrongly) perceive as affecting our survival and reproduction.

## **Weakness of will**

Most human beings suffer difficulties identifying what they want and being consistent about it. This may be a natural consequence of at least two factors: conflict between mental modules and maladapted discounting.

The modularity internal to the human mind causes a typical trade-off between specialization gains and transaction costs. Developing specialized mental mechanisms to solve different problems might incur substantial transaction costs that will take the form of discrepancies and conflicts between these specialized modules, as each one optimizes the allocation of scarce resources toward a different goal. In addition, it would be too costly to eliminate these transaction costs, achieving a perfect fit between the aggregation of locally optimal partial solutions reached by different modules and the optimum for the whole. Current knowledge on consciousness is too weak to reach a definite conclusion,<sup>4</sup> but the existence of internal mental conflict is supported by biological evidence on apparently awkward conflicts, as those between

---

<sup>4</sup> For some, it is clear that no part of the brain decides, as mental “supervisors” fail systematically and suffer self-deception. Self control is therefore claimed to be merely a sort of “spin doctor,” an illusion (Pinker, 2002, 42-43). For a more nuanced view, see Ramachandran (2004).

cells and organs within a body as well as those between a pregnant mother and the fetus in her womb.

Daily life is also full of instances in which we make inconsistent decisions over time, from the difficulties of following a diet or quitting smoking, to the tricks we use to get out of bed in the morning or to study regularly. These weaknesses of will may emerge because of maladaptation to our current environment, which is substantially less risky than our ancestral environment and therefore makes it optimal to postpone gratification whilst we have evolved to emphasize present consumption. Let us see why.

Human beings constantly allocate scarce resources over time through saving and investment decisions which confront present against future consumption, decisions which are driven by both reason and emotions. Evolution has also developed automatic mechanisms to cope with this allocation problem and to maximize the chances of reproduction. It is also likely that innate traits have evolved for decision making, a sort of “subjective discount rate,” finely tuned to our expected longevity and the level of risk in our environment, and both affecting and embedded in our emotions.<sup>5</sup>

It is likely that such innate discount rate is too high for our current stable environment and long life. Risks were much greater in the ancestral past, because of lesser control on nature and the prevalence of warfare. Therefore, life expectancy was very short and, in accordance with such circumstances, we probably evolved a high subjective discount rate which ceased to be optimal when, quite recently, we achieved a safer environment.

This may explain that we now need artificial technologies of self-control to be able to postpone gratification and better adapt to our environment. For instance, it is clear that much of our educational effort is directed to change children’s preferences in order to postpone gratification, inculcating them with a lower discount rate. This cultural lowering of the subjective discount rate is most noticeable when lacking: many children raised in broken families and ghettos easily fall prey of drug addiction and all kind of short-sighted behavior.

---

<sup>5</sup> In fact, the brain’s response to short term opportunities is mostly emotional, taking place in the limbic system, while long term rewards are governed by reason and calculation, triggering brain activity in the prefrontal cortex (McClure *et al.*, 2004).

## **V. Cooperation**

Specialization increases productivity but requires cooperation, and is often costly to achieve. Only in the less-interesting cases do cooperation benefits come at no cost to cooperators. Symbiotic interaction is the paradigm of this sort of costless and non-conflictive cooperation. It explains, for example, why some animals live together in amorphous shoals or herds in order to be better protected from predators. When we trade simple commodities in the spot market we are also close to such ideal of mutuality.

The most interesting cases of cooperation, however, are those in which cooperation benefits involve substantial costs for cooperating parties. Their interaction is therefore prone to conflicts of interests, as each cooperator tries to reap the benefits of cooperation without paying the corresponding share of the costs. In such cases, cooperation requires enforcement mechanisms to make sure that parties comply.

This enforcement is produced by different means, which rely more or less heavily on innate instincts. The most simple of these instincts are linked to genetic relatedness, which grounds cooperation between relatives. The most complex are those instincts supporting cooperation between genetically unrelated individuals. They play a role even for achieving cooperation between total strangers.

Let us examine some elements of this arsenal of cooperative instincts, how they work and to what extent they are maladapted.

### **V.A. Instinctive cooperation**

#### **Genetic relatedness**

By helping their children, parents promote the survival of their own genes. More precisely, genes driving parents to help their children had a better chance of survival and became dominant. This explains why humans in all cultures are benevolent toward their own descendants and

relatives, the more so the greater their genetic relatedness, leading to nepotism, which has been shown to be universal.<sup>6</sup>

The common practice of taking the family as a rhetorical model when we want to emphasize cooperation suggests that genetic relatedness is effective. Managers, for example, claim that “our firm is a family,” and believers of many faiths treat co-believers as brothers and sisters, and priest as fathers. In addition, genetic relatedness does not require external enforcement, because parties are preprogrammed to cooperate. On the negative side, however, it motivates rent seeking in the form of cuckoldry and correspondingly, spending resources to avoid it. Furthermore, nepotism often conflicts with “higher” forms of cooperation, which explains that, despite being universally present, the most developed cultures try to repress it. In addition, cooperation grounded on genetic relatedness is limited to a few individuals, and, as a consequence, it does not allow much specialization.

### **Emotional commitment**

Genetic relatedness is only the tip of the iceberg of cooperative instincts. Even strangers playing standard non-repetitive cooperation games cooperate more than pure logic predicts, especially when they are allowed to talk with the other players (for example, Valley *et al.* 2002). It seems that we are able to evaluate, detect, or link emotionally to our mates, and this allows us to overcome part of the cooperative dilemma we are facing.

The implementation of these detection and reciprocity strategies requires a sophisticated mind, which has to be capable of, at least, forming cooperative initial expectations, foreseeing future interactions, distinguishing cheating from cooperative behavior by partners and keeping a record of past interactions. Human minds are equipped with tools designed for overcoming these problems because they were important in our evolutionary past.

The simplest evidence on the existence of these instincts comes from the physiological consequences of insincerity: blushing often follows lying, and most people cannot avoid showing their feelings. This explains why business travel continues being important in these times of the

---

<sup>6</sup> These ideas of “kin selection” or “inclusive fitness” were developed by Williams and Williams (1957); Hamilton (1963, 1964); and Maynard Smith (1964). Brown (1991) points out the universal presence of nepotism.

internet and teleconference, as people have difficulty in evaluating trustworthiness by telephone or e-mail.

These cooperative tools, from the relative simple, such as facial expressions, to the most sophisticated, such as love, are instinctive and not calculative because it would be inefficient to solve most of these cooperation problems by rational calculation, using a general-purpose mental process. Couples rely on love and attachment to safeguard their cooperation. In a similar way, criminals ground their cooperation on their urge to defend their reputation and territory, which leads them to costly and seemingly irrational fights and revenges. Emotions thus provide solutions that “better than rational” when they commit individuals to a certain behavior which is optimal in the long term.

Emotional responses often seem irrational and maladaptive, as when we die to save a loved one or to punish an enemy. However, these emotions are part of an efficient commitment strategy. If he is willing to die for her, she will be much more willing to take him as a partner. If a criminal retaliates all offenses, he will deter potential offenders. The problem of both, lovers and criminals, is to make their disposition credible, as, after the fact, it would pay to change their minds and avoid his giving up his life or inflicting a costly punishment. Being emotionally driven provides this credibility. He falls in love so much that when faced with the situation of risking his life he does not calculate costs and benefits. He simply cannot help but throwing himself to save her. The criminal’s rage and urge for revenge may play a similar role. For both, their emotional responses will occasionally seem inefficient when they are activated, but this apparent inefficiency is hiding that such instances of activation make it possible achieving greater efficiency in many other cases. In a sense, ex post “irrational” emotions are introducing greater ex ante rationality.

### **The tools of reciprocity**

The human brain is also well endowed to distinguish cooperators from cheats among potential partners and to distinguish cooperative from cheating behavior. Think that, to be fruitful, emotional commitment requires that prospective parties are able to distinguish cooperative individuals from cheats beforehand. Similarly, reciprocity, which is probably the

basis of most cooperation in modern societies,<sup>7</sup> also requires that participants be able to distinguish at least compliant from cheating behavior after the transaction. For instance, playing even the simplest reciprocity strategy, such as “tit-for-tat,”<sup>8</sup> requires us to distinguish cooperative from cheating behavior a posteriori.

Both abilities, detecting cheats *ex ante* and *ex post*, are related and have to cope with the possibility of error and mimicry: those who read a cooperative move as cheating when playing tit for tat are inviting retaliation on themselves; those who take a cheat for a cooperator will get the worst of their association. Understandably, human beings seem to have developed specialized innate abilities to detect cheating behavior, as well as to signal and distinguish cooperators, which make it possible for human populations to reach stable polymorphic equilibria with different types—for instance, cooperators, reciprocators and cheats.<sup>9</sup> Some biologists have even argued that our brains develop in the “arms’ race” of deception and detection. In sum, reciprocity is grounded on complex mental tools of detection and commitment. Two of these tools are our ability to detect cheats and our urge to retaliate when we feel that we have been cheated.

### **Cheating detectors**

The presence of mental resources specialized in detecting cheats has been shown by an adaptation of the Wason psychological tests by Leda Cosmides (1985, 1989; Cosmides and Tooby, 1992). The original test measures humans’ logic ability by trying to find out how good we are at falsifying hypotheses. For example, a set of four cards with letters on one side and numbers on the other, such as

D

F

3

7

---

<sup>7</sup> Main ideas on what is often labeled “reciprocal altruism” were developed by Williams (1966), Trivers (1971, 1985) and Alexander (1987). Reciprocity has been shown to be a human universal, with similar results being obtained in experiments run in different cultures and greater cooperation observed in societies in which people rely more on market exchanges in their daily lives (Henrich *et al.*, 2001, 2005).

<sup>8</sup> Tit-for-tat consists of cooperating in the initial round and replicating in other rounds the conduct of the other player in the preceding round.

<sup>9</sup> See Maynard Smith y Price (1973) and Frank (1987), and, for a recent empirical test, Kurzban and House (2005).

is shown to a sample of individuals who are then said to test the rule “If a card has a *D* on one side, it must have a 3 on the other” by turning over as fewer cards as possible. It results that only between 5% and 15% of people get it right (*D* and 7 in this example).<sup>10</sup>

However, as shown by Cosmides, the same task becomes much easier if it is expressed in terms of contractual exchange, when finding a false is equivalent to detecting a cheat. Imagine, for example, that you are enforcing the rule “If a person is drinking beer, he must be 18 or older” by checking either their drink or their age.

*Beer drinker*      *Coke drinker*      *25 year old*      *16 year old*

In this case most people get it right (checking the beer drinker and the 16 year old) despite the fact that the logical structure of the problem is exactly the same than before. Furthermore, the improvement in solving the puzzle is not caused by the concreteness of the story because most people also fail when it is told in a concrete setup without a cheating element. For example, falsifying the rule “If a person eats hot chili peppers (HCP), then he must drink cold beer” where SCP represents sweet chili peppers, is not easier than the example with the abstract cards *DF37*.

*Eats HCP*      *Eats SCP*      *Drinks beer*      *Drinks Coke*

In conclusion, it seems that our ability to obtain the right solution is higher in a cheating situation thanks to our use of mental resources which work faster and better than when processing the abstract rules of logic.

### **Strong reciprocity**

Reciprocity seems well suited for repeated interaction but experiments also show that humans often practice a *strong* form of reciprocity that is well suited to one-shot transactions: we are willing to incur costs to punish cheats even when there is no prospect of further interaction.

---

<sup>10</sup> The card with a *D* is informative, because if there was not a 3 the rule would be falsified. The card with a 7 is also informative, because if there was a *D* the rule would be falsified. The card with an *F* is not informative, because, whatever the number on the other side, it would comply with the rule. The card with a 3 is not informative either because the rule does not forbid having a 3 and any other letter.

Interestingly, this propensity to punish ends up achieving greater cooperation, when parties anticipate the possibility of costly retaliation.

Both results have been proved in many experiments (Fehr and Gächter, 2000). In those with an “ultimatum” game, an individual, *A*, divides 1000 € between himself and *B*, but none of them gets a cent if *B* rejects the offer. Usually, the distributor *A* divides by half and, interestingly, *B* rejects offers below 30%, even for stakes of as much as three months’ earnings. Given that distributors are less generous when *B* cannot reject and *A* acts as a dictator, it seems that the expectation of *B*’s rejection helps in eliciting generous offers.

The psychology of retaliation is also revealed in experiments that test our reactions in “public good” games. In these, a number of people contribute money to a common pool, expecting to be compensated later with an equal share in a multiple of the pool. In one-stage games, people often contribute half their wealth. In multiple stages, however, people start contributing more but their contributions decay with time and approach zero at the end. This fall is not driven by learning but likely by the fact that, in the experiment, the only punishment cooperators can inflict on free riders is by cheating themselves. Remarkably, when the game is redefined so that cooperators can punish free riding even at a cost and without prospect of future interaction (“strong reciprocity”), they do punish them, and this increases cooperation. Therefore, depending on circumstances, either cheats lead cooperators who are incapable of retaliation to cheat, or cooperators who are willing to incur costly retaliation lead cheats to cooperate.

## **V.B. Ecological cooperation: the maladaptation of our instincts**

Whatever the power of cooperative instincts, their adaptation to the hunter-gatherer environment of our ancestors means that they may be maladapted to the cooperative demands of our current environment. We will now comment on the limits of cooperation grounded on cooperative instincts to examine, next, how these limits constrain the characteristic form of modern cooperation—that taking place through market exchange.

### **The limits of instinctive cooperation**

Cooperative instincts are powerful but limited to certain forms of cooperation, mainly within small groups of known people. These limits are most prominent for genetic relatedness, which



promotes cooperation only among relatives, but other mechanisms have also their particular, even if less strict, limits. Thus, most emotional commitment and cheating detectors require personal interaction. Direct reciprocity is also limited to relatively small groups because it requires us to know others and to keep track of their behavior.

We now live in large groups, with anonymous and more impersonal, indirect and superficial interactions. In part, we rely on direct reciprocity. For example, brand managers are well aware that consumers have a personal and emotionally loaded relationship with the brands that they consume. We rely more, however, on mechanisms of indirect reciprocity. In the market, for example, we reward a merchant who rewards another merchant and so on, until, after several more steps, a manufacturer is rewarded for producing a good product. But not only in the market place. The legal order is grounded on indirect reciprocity by the use of third party (mainly, judicial) enforcement.

### **The unnaturalness of market exchange**

Most of these mechanisms of indirect reciprocity are institutional. They are designed to promote a certain type of exchange for which we are poorly endowed by nature, for example trade between anonymous parties. The design and difficulties of such institutions are often connected to this intrinsic maladaptation.

This reasoning is especially applicable to markets. Market dealings may suffer all sorts of problems and therefore require substantial institutional support. Given that, in the evolutionary timescale, at least some forms of market exchange are very new, they are prone to conflict with our instincts.<sup>11</sup> Several types of conflict appear when considering that these instincts, well adapted to the economic environment of our hunter-gatherer ancestors, will tend to bias us against anonymous parties and at least certain forms of trade, insurance and capital, including wealth accumulation and credit:

---

<sup>11</sup> This argument provides a common and more solid ground to the pioneering and rival arguments by Polanyi (1944) on the limits of market-type relations and the resistance of societies to the dominance of such relations; and Hayek (1944) on the opposing rules of the “extended order of cooperation through markets” and the more intimate and personal order. The danger that the primitive collectivistic leanings of human beings pose to the market has also been stressed by Smith (2003) from the perspective of experimental economics.

Our hunter gatherer grandparents lived in small bands of no more than 100 or 150 individuals, which limited social interaction and economic specialization, as most interactions were with identifiable people who you know personally. A bias against unknown, anonymous, people likely developed as a consequence<sup>12</sup>. In addition, the value added by those making indirect trade possible (intermediaries) and abstract forms of trade (e.g., in services, intangibles or capital) may be more difficult to grasp. Furthermore, warfare among bands was prevalent, much more so than in modern societies (Keeley, 1996), which throws doubt as to how much are we naturally endowed to gain through production and trade or, instead, to expropriate strangers through violence.

With respect to distribution, hunter-gatherers distributed their production following a mixed pattern of sharing and private appropriation. This agreed with economic logic, as they shared resources (big game) as insurance against exogenous risks and privately appropriated those resources (fruits) whose production would have suffered the most from the perverse incentives caused by sharing (Cashdan, 1980; Kaplan and Hill, 1985; Bailey, 1992). This predisposition to compensate exogenous risks now likely poses systematic problems to insurance markets. Given that human societies are predisposed to compensate bad luck ex post, it does not make sense to buy insurance ex ante. The argument can be applied widely in all sorts of insurance, from farm production to earthquakes or, most importantly, healthcare; and also provides support for welfare states.

Lastly, the ancestral situation with respect to capital and technology also holds important consequences. First, the need of mobility meant that our ancestors only accumulated portable capital. This may have hindered our ability to understand the productivity of capital and the basis for paying interest. Second, technical change was extremely slow, causing a practical absence of economic growth (Kremer, 1993). This may have predisposed humans against inequality and even wealth because, in the absence of growth, the economy becomes a zero-sum game and individual inequality and wealth are more likely to proceed from expropriation than from socially productive activities.

---

<sup>12</sup> Excessive attention therefore may be paid now to identifiable individuals, to the detriment of anonymous parties by both the political process (Rubin, 2002, 153-181) and judges (Arruñada and Andonova, forthcoming).

## **VI. The Role of Institutions**

Against this background, institutions act as rationalizing and cooperative mechanisms that enhance our fitness in new environments. They are, however, grounded on our human nature: in a sense they “recruit” and mold our instincts to build more effective mechanisms. We will briefly examine this recruitment process, to focus on the role of institutions.

### **VI.A. Instincts as building blocks of institutions**

Natural selection has often recruited body organs to perform functions different from the ones that they were originally designed to do. For example, our limbs were developed starting from the swimming fins of our fish ancestors. Similarly, institutions recruit instincts as building blocks of their machinery, often to create enforcement mechanisms. A simple case is that of disgust, an emotion that was originally useful to avoid poisoning, an important risk for omnivorous animals. For example, food taboos (e.g., against eating pork) seem to be exploiting the psychology of disgust during the period when children learn their food preferences, probably to make for them more difficult to interact with members of neighboring groups when they are grown up. These cultural taboos show enormous variety but all of them rely on the same instinctive mechanisms. In the case of disgust, this mechanism is related to the idea of a polluting substance, what explains that the feeling of disgust is independent of the amount of contact or how much the substance is diluted.

Applications of the recruitment of instincts for higher ends abound. We have seen above how our drive for fairness triggers retaliation *ex post* and elicits cooperation *ex ante*. Many religions rely on fear of a punishing God and some on love of God to motivate good deeds, and feelings of shame and guilt are present in most correctional institutions. In general, different emotions seem to be active in different kinds of enforcement mechanisms.

### **VI.B. Institutions as complements of human nature**

Filling the adaptation gap between our ancestral and current environments requires us to manage our instincts on both the rationality and cooperation fronts. In terms of rationality, the

paramount issue is one of self-control, postponing gratification, while in terms of cooperation, it is fundamental to control antisocial behavior. Self-control means greater control of our emotions to improve our individual fitness; for example, instilling a lower subjective discount rate by education. Social control means controlling free riding. In a sense, it can be understood as a way of dealing only with cooperators.

### **Institutions as enforcement**

Focusing on social control, institutions act as enforcement mechanisms which permit human groups to achieve greater cooperation within the group and makes them more competitive against other groups. Considering which of the exchange's parties is acting as enforcer, three kinds of enforcement may be distinguished, and institutions play an important role in all of them.

Under *first party enforcement*, it is the obliged party who acts as enforcer, relying for punishment on emotions like guilt and shame. To function properly, it requires previous indoctrination and selection of types before contracting. The role of institutions lies mainly in defining and indoctrinating a moral code, whose violation triggers innate guilt and shame emotions. On occasion, institutions are also involved in helping to enforce the moral code.<sup>13</sup> In addition, this code may have very different properties and, therefore, facilitate different kinds of cooperation. For instance, Protestantism seems to promote values that support anonymous exchange while Catholicism is more hospitable to smaller communities (Arruñada, 2004).

*Second party enforcement* is grounded on reciprocity, as the receiving party is the one who sanctions the defaults. Emotions triggering seemingly inefficient ex post retaliation act as important enforcement mechanisms, deterring cheating in anticipation of retaliation. A key activity is the correct evaluation of performance to avoid unjustified retaliation. Understandably, evolution has dedicated specialized mental resources to provide us with innate cheating detectors. The role of modern institutions, however, is often to channel or impede private retaliation. For example, criminal law punishes retaliation and rules and courts restrain asymmetric relational contracting.

---

<sup>13</sup> Sometimes, institutions also help the individual to enforce the code, as it happens in Catholic oral confession (Arruñada, 2007).

Within *third party enforcement*, other persons act as enforcers. It can be informal and decentralized, as in the functioning of a commercial market or a social network, or highly formal and centralized, as in judicial procedures. Decentralized enforcement relies on different forms of reputational investments and gossip, including at present the activity of the mass media. Centralized enforcement relies on litigation and accumulation of sentences. It suffers the same problem as any other specialization: positive transaction costs, given that third party enforcers, as any other specialist, may pursue their own interests to the detriment of the underlying transactions.

## **VII. Where are we humans heading?**

Analyses of human behavior that point out the presence of innate traits used to be wrongly read as genetic determinism, as prevalence of nature over nurture. This should not happen with modern cognition science, because it surpass this controversy on the relative importance of nature and nurture on behavior by emphasizing that both, nature and nurture, interact in a way that makes them not separable. They act as complements more than as substitutes. For example, children learn to speak different languages by growing up in different environments, but they learn by using a highly developed innate learning mechanism that includes most structural grammar, what explains why they learn to speak so fast and suddenly, almost exploding to speak between two and three years of age (Pinker, 1994; Baker, 2001).

The interaction also takes place at the institutional level. The previous analysis of the adaptive role of institutions should help us in avoiding the mistake of genetic determinism because institutions interact with our instinctive traits, both recruiting them and complementing them. Institutions mold the nurturing process and display a full array of enforcement mechanisms that greatly affect our self and social control abilities. And institutions are the product of intentional design, relying on instincts, as explained above, but intentionally designed. This might reduce the influence of genetic selection, sitting human beings at the wheel of their destiny.

We now interact technologically with our nature, as we have greater control over it: we achieve the pleasures that nature used to drive our behavior without incurring the costs of such

behaviors: contraceptives allow us to have sex without producing children; saccharine satisfies our sweet teeth without obesity; and so on.

More importantly, institutions allow greater human interaction, enhancing specialization and multiplying our productivity. Institutional enforcement not only boosts within-group cooperation by punishing free riders but also enlarges the cooperative group. For instance, the rule of law makes trade with strangers much easier. It also channels between-group conflict to productive ends, by precluding violence, as it happens, for instance, in market competition between firms, ending up with multiple levels of cooperative groups.

Furthermore, the process of institutional change is different from natural selection. It is constrained by the genetic background, but is substantially influenced by learning, decision-making and imitation. Interbreeding is also possible, triggering processes similar to contagion and infection. To some extent, acquired behavioral features may also be transmitted from one generation to the next.<sup>14</sup>

Institutional change is also *intentional*, the consequence of individual decisions.<sup>15</sup> Intentional does not necessarily mean successful, however. We develop technologies only to be surprised by their unintended consequences, and our abilities at institutional design are probably even lower. Therefore, even if we are now at the wheel of our destiny, we are barely learning to drive it.

---

<sup>14</sup> It is therefore, Lamarckian. Jean-Baptiste Lamarck argued at the beginning of the nineteenth century that traits acquired by an organism on to its progeny. For example, the long neck of the giraffe would result from generations of animals stretching to reach the highest leaves.

<sup>15</sup> Analyzing how are institutions designed would exceed the proper scope of this work, as it would make necessary going into the many theories of cultural evolution. See, for instance, Dawkins (1976, 1982), Boyd and Richerson (1985) and Sperber (1996).

## VIII. References

- ALEXANDER, Richard D. 1987. *The Biology of Moral Systems*, Aldine de Gruyter: Hawthorne.
- ARRUÑADA, Benito. 2007. "Third Party Moral Enforcement: The Rise and Decline of Christian Confession," Universitat Pompeu Fabra, Economics and Business Working Paper Series 653, January, <http://www.econ.upf.es/cgi-bin/onepaper?653> (accessed January 13, 2005).
- ARRUÑADA, Benito. 2004. "The Economic Effects of Christian Moralities," Universitat Pompeu Fabra, Economics and Business Working Paper Series 743, March, <http://www.econ.upf.es/cgi-bin/onepaper?743> (accessed January 13, 2005).
- ARRUÑADA, Benito, and Veneta ANDONOVA. Forthcoming. "Judges' Cognition and Market Order," *Review of Law and Economics*.
- BAILEY, Martin J. 1992. "Approximate Optimality of Aboriginal Property Rights," *Journal of Law and Economics*, 35, 183-98.
- BAKER, Mark. 2001. *The Atoms of Language*, Basic Books, New York.
- BOYD, Robert, and Peter J. RICHERSON. 1985. *Culture and the Evolutionary Process*, University of Chicago Press: Chicago.
- BROWN, Donald. E. 1991. *Human Universals*, McGraw-Hill: New York.
- CASHDAN, Elizabeth A. 1980. "Egalitarianism among Hunters and Gatherers," *American Anthropologist*, 82, 116-20.
- COASE, Ronald H. 1976. "Adam Smith View of Man," *Journal of Law and Economics*, 19, 529-46.
- COSMIDES, Leda, and John TOOBY. 1994. "Better than Rational: Evolutionary Psychology and the Invisible Hand," *American Economic Review*, Papers and Proceedings, 84, 327-32.
- COSMIDES, Leda, and John TOOBY. 1992. "Cognitive Adaptations for Social Exchange," in *The Adapted Mind*. J. H. Barkow, L. Cosmides and J. Tooby, eds. Oxford University Press: New York, 163-228.
- COSMIDES, Leda. 1985. *Deduction or Darwinian Algorithms? An Explanation of the 'Elusive' Content Effect of the Wason Selection Task*, Ph.D. Dissertation, Department of Psychology, Harvard University.
- COSMIDES, Leda. 1989. "The Logic of Social Exchange: Has Natural Selection Shaped How Humans Reason? Studies with the Wason Selection Task," *Cognition*, 31, 187-276.
- DAMASIO, Antonio R. 1994. *Descartes' Error: Emotion, Reason and the Human Brain*, Grosset/Putnam: New York.
- DAWKINS, Richard. 1976. *The Selfish Gene*, Oxford University Press: Oxford.
- DAWKINS, Richard. 1982. *The Extended Phenotype*, Oxford University Press: Oxford.
- FEHR, Ernest, and Simon GÄCHTER. 2000. "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives*, 14, 159-81.
- FRANK, Robert H. 1987. "If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience?," *American Economic Review*, 77, 593-604.

- GIGERENZER, Gerd. 2000. *Adaptive Thinking: Rationality in the Real World*, Oxford University Press: New York.
- GIGERENZER, Gerd, and Peter M. TODD. 1999. "Fast and Frugal Heuristics: The Adaptive Toolbox," in *Simple Heuristics that Make Us Smart*. G. Gigerenzer, P. M. Todd, and the ABC Research Group. Oxford University Press: New York, 3-34.
- HAMILTON, William D. 1963. "The Evolution of Altruistic Behavior," *American Naturalist*, 97, 354-56.
- HAMILTON, William D. 1964. "The Genetical Evolution of Social Behaviour I and II," *Journal of Theoretical Biology*, 7, 1-52.
- HAYEK, Friedrich A. 1944. *The Road to Serfdom*, University of Chicago Press, Chicago.
- HENRICH, Joseph, Robert BOYD, Samuel BOWLES, Colin CAMERER, Ernst FEHR, Herbert GINTIS, Richard MCELREATH, Michael ALVARD, Abigail BARR, Jean ENSMINGER, Kim HILL, Francisco GIL-WHITE, Michael GURVEN, Frank MARLOWE, John Q. PATTON, Natalie SMITH, and David TRACER. 2005. "'Economic Man' in Cross-cultural Perspective: Behavioral Experiments in 15 Small-scale Societies," *Behavioral and Brain Science*, 28(6), 795-815.
- HENRICH, Joseph, Robert BOYD, Samuel BOWLES, Colin CAMERER, Ernst FEHR, Herbert GINTIS and Richard MCELREATH. 2001. "Cooperation, Reciprocity and Punishment in Fifteen Small-scale Societies," *American Economic Review*, 91, 73-78.
- HODGSON, Geoffrey M. 2004a. *The Evolution of Institutional Economics: Agency, Structure and Darwinism in American Institutionalism*, Routledge: London and New York.
- HODGSON, Geoffrey M. 2004b. "Veblen and Darwinism," *International Review of Sociology/Revue Internationale de Sociologie*, 14, 343-61.
- KAPLAN, Hillard, and Kim HILL. 1985. "Food Sharing among Ache Foragers: Test of Explanatory Hypotheses," *Current Anthropology*, 26, 223-39.
- KEELEY, Lawrence H. 1996. *War Before Civilization: The Myth of the Peaceful Savage*, Oxford University Press: New York.
- KLEITHER, Gernot D., Marianne KREBS, Micheal E. DOHERTY, Hugh GARAVAN, Randall CHADWICK and Gregory BRAKE. 1997. "Do Subjects Understand Base Rates," *Organizational Behavior and Human Decision Processes*, 72, 25-61.
- KREMER, Michael. 1993. "Population Growth and Technological Change: One Million B.C. to 1990," *Quarterly Journal of Economics*, 108, 681-716.
- KURZBAN, Robert and Daniel HOUSER. 2005. "Experiments Investigating Cooperative Types in Humans: A Complement to Evolutionary Theory and Simulations," *Proceedings of the National Academy of Sciences*, 102, 1803-807.
- MARKOCZY, Livia, and Jeff GOLDBERG. 1998. "Management, Organization and Human Nature: An Introduction," *Managerial & Decision Economics*, 19, 387-409.
- MAYNARD SMITH, John. 1964. "Group Selection and Kin Selection," *Nature*, 201, 1145-47.
- MAYNARD SMITH, John, and G. R. PRICE. 1973. "The Logic of Animal Conflicts," *Nature*, 246, 15-18.
- McCLURE, Samuel M., David I. LAIBSON, George LOEWENSTEIN, and Jonathan D. COHEN. 2004. "Separate Neural Systems Value Immediate and Delayed Monetary Rewards," *Science*, 306, 503-507.
- PINKER, Steven. 1994. *The Language Instinct: How the Mind Creates Language*, Morrow: New York.
- PINKER, Steven. 2002. *The Blank Slate: The Modern Denial of Human Nature*, Viking: New York.



- POLANYI, Karl. 1944. *The Great Transformation*, Bacon Press: Boston.
- RAMACHANDRAN, V. S. 2004. *A Brief Tour of Human Consciousness. From Impostor Poodles to Purple Numbers*, PI Press: New York.
- RUBIN, Paul H. 2002. *Darwinian Politics: The Evolutionary Origin of Freedom*, Rutgers University Press: New Brunswick.
- SIMON, Herbert A. 1956. "Rational Choice and the Structure of Environments," *Psychological Review*, 63, 129-38.
- SIMON, Herbert A. 1962. "The Architecture of Complexity," *Proceedings of the American Philosophical Society*, 106, 467-482.
- SMITH Adam. 1759. *The Theory of Moral Sentiments*, London: A. Millar, Edinburgh: A. Kincaid & J. Bell.
- SMITH, Vernon L. 2003. "Constructivist and Ecological Rationality in Economics," *American Economic Review*, 93, 465-508.
- SPERBER, Dan. 1996. *Explaining Culture: A Naturalistic Approach*, Blackwell: London.
- TOOBY, John, and Irven DeVORE. 1987. "The Reconstruction of Hominid Behavioral Evolution Through Strategic Modeling," in *Primate Models of Hominid Behavior*. W. G. Kinzey, ed. SUNY Press: Albany, 183-237.
- TOOBY, John, and Leda COSMIDES. 1992. "The Psychological Foundations of Culture," in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Jerome H. Barkow, Leda Cosmides and John Tooby, eds. Oxford University Press: New York, 19-136.
- TRIVERS, Robert L. 1971. "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology*, 46, 35-57.
- TRIVERS, Robert L. 1985. *Social Evolution*, Benjamin/Cummings: Menlo Park.
- VALLEY, Kathleen, Leigh THOMPSON, Robert S. GIBBONS and Max H. BAZERMAN. 2002. "How Communication Improves Efficiency in Bargaining Games," *Games and Economic Behavior*, 38, 127-55.
- WILLIAMS, George C., and Doris C. WILLIAMS. 1957. "Natural Selection of Individually Harmful Social Adaptations among Sibs with Special Reference to Social Insects," *Evolution*, 11, 32-39.
- WILLIAMS, George Christopher. 1966. *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*, Princeton University Press: Princeton.