

Miguel Angel Latouche R

Goethe University Frankfurt

Central University of Venezuela

Institutions and the Problem of Social Order.¹ A Normative Approximation²

*It is the success and failures in human organizations that account
for the progress and retrogression of societies*

Douglass North

People must learn to play games in an imperfect world

Binmore and Samuelson

Abstract: This paper discusses the effect of social structure in the development of cooperative behaviour. We will argue that structural conditions shape the scope of social interactions. Strategic exchanges among the members of society are built in the context of the parameters that are defined by the institutional framework in which social life develops. Cooperation is not the result of natural development, nor does it depend on altruistic behaviour; instead, it results both from the configuration of normative mechanisms of control and incentives systems, and from the evolution of moral foundations resulting from learning processes. We will use the Stag Hunt Metaphor to argue about the impact that context may have in determining the scope of human interactions and about the importance of defining moral mechanisms that favour the construction of a stable social order.

Keywords: evolution, social contract, normativity, interest, equilibrium

Introduction

This inquiry builds on a large academic tradition: that according to which the Social Contract is the fundamental building block of social interactions. Social order requires the development of mechanisms that will regulate the scope of in-

¹ This article was written during a postdoctoral stay at the Cluster of Normative Orders, Goethe University, Frankfurt.

² To Professor Alfredo Barragan in memoriam.

teractions and exchanges that normally result from our dealings, interactions and associations with other human beings. From the moment we are born, we learn about the *rules of the game* that actors in our society have determined as valid mechanisms for the regulation of behaviour and economic exchange. Such rules constitute the normative structure of society. Analytically, these rules allow us to determine the payoffs that those associated would receive on specific social situations, and empirically they help us to work out the problems of building trust and developing cooperative behaviour. Game theory has a series of models and metaphors that assist us in thinking about the rank of practical solutions that would be produced by a given social design. Each of these metaphors proposes an outcome that results from the interactions of the autonomous players involved in the model. We will work around the idea that the institutional framework will determine the scope of equilibrium solutions to human interactions. We will use the Stag Hunt Game, as a metaphor for cooperative N-Players social games; from there we will discuss the normative impact of a given equilibrium on the structure of political associations.

The study of institutions from a normative perspective allows for the development of a method for the analysis of democracy based on the definition of a justificatory structure that validates social cooperation under valid criteria of distributive justice. We assume that the contemporary crisis of democratic legitimacy is in response to failures in the normative capacity of the political system to correctly regulate interactions and exchange and to distribute social goods and justice among the population. As the fundamental components of the democratic social contract erode and the levels of satisfaction and support are reduced, populism and authoritarianism start being perceived for many as valid alternatives. Clearly, there are circumstances in which institutional designs do not work properly. In such cases, society's capabilities to regulate interactions will be reduced and social order will become unstable. This instability reduces guarantees for a fair redistribution of costs and benefits, associated with social interaction and exchange, and increases the costs of cooperative behaviour, all of which will, most likely, produce social unrest and political instability.

We are assuming that democracy is a cooperative enterprise that, when it works correctly, has a positive impact on the well-being of the population. As such, it needs to guarantee an increasing function of social utility. Democracy must allow people to act freely and to increase their individual utility within the limits established by rules of the game that oblige them to show respect for the rights of others, observe their own responsibilities and to accomplish their duties. When democracy produces valuable results, individuals tend to become (more) interested in protecting and reproducing social order, after all, it will benefit their own interests, due to the expectancy of a future distribution of the

“cooperative surplus”. A fundamental problem faced by ‘social designers’, then, would be to determine the most appropriate architecture for the institutional framework. Conflicts of interests should be solved in the context of a functioning democracy and not by imposition. One would understand that autonomous rational individuals will be capable of reaching social agreements that will benefit their collective interests and of developing mechanisms for its enforcement. There is, then, a requirement for civilized solutions for the problems related to social interactions and cohabitation which, at the same time, are associated with the evolution of a democratic social order. We consider that a democratic society must provide a normative framework that makes cooperative strategic interactions possible in the long-run. We will argue that it would only be possible under conventions normally accepted and broadly validated by the members of society as a fair mechanism for the regulation of behaviour that results from a transaction process. We next present a discussion regarding the role of institutions in the development of social order.

Social dilemmas

Our life develops in the context of complex social interactions. Each one of us ends up being part, since we were born, of the particular social arrangement in which our lives develop in association with others (Rawls, 2004). The complexities of human coexistence rest on building solutions to problems associated with the distribution of costs and benefits, the production for consumption, and the implementation of mechanisms for making viable collective associations. From a contractual perspective, one would hope society uses its resources to: guarantee the reduction of social conflict; moderate collective action problems; diminish delinquency and social exclusion; facilitate productive activities; ease the hardship of poverty; create well-being; increase economic growth and prosperity; secure the futures of coming generations; strengthen democratic values and standards; and, so on. As we all know, however, that is not always the case. It can be said that, above all, society is shaped by collective arrangements intended to satisfy particular and collective preferences, related to belief systems, values, and needs; this is so, in order to fulfil the aggregation of individual demands in a situation of resources scarcity, and to guarantee the functionality of the supply side.³ The design of an institutional framework that provides for the satisfaction of individual needs and the functionality of social interaction and exchange is, then, the building stone of a social struc-

³ We are referring here to the conditions that allow individuals to act publicly favouring the construction of a cooperative surplus: freedoms, liberties, opportunities, security.

ture based of a functional equilibrium.⁴ After all, people expect to live a life worth living, that is, a life that allows for the satisfaction of their needs and the highest possible fulfilment of their hopes and expectations (Sen, 2000), one that guarantees the materialization of individual projects in terms of their own expectations, within the context of the society in which they were born (Rawls, 2004).

Each social group would come across its own solutions to the problems it faces in the ambit of its own historical development. Collective agreements would be a necessary condition for the functionality of political associations. Those agreements are constructed within the context of the particularities of the social and political development of a given society. As society develops, it incorporates the particularities of the population in terms of culture, values, identities, aspirations and preferences. At the same time, these agreements must be structured in terms of resources availability, technological bulk, productive capacities and distributive capabilities (Binmore, 1989). In that sense, we cannot expect uniformity in the shape that different social arrangements will adopt in one society or in another. People's expectations and preferences will differ because of the particularities of the historical path of their own societies.⁵ In fact, one can expect that between one social arrangement and another both structure and efficiency levels will vary (North, 1990). Context, after all, matters (Goodin, Tully, 2006).

An understanding of social ethos requires an evaluation of the particular characteristics of a specific society. Even assuming that people tend to maximize its *collective welfare function*, we must be careful with our understanding of social processes. We should not study the problem of expected utility maximization in strict economic terms. Instead, flexibility is required, as we need to understand well-being beyond the issues of wealth production, economic growth, or income distribution. That is, we need to deal with the collective welfare function as a socio-political category. The process of building-up social arrangements is not neutral, but, rather, is culturally biased. After all, culture, values, attitudes, preferences, and actors' capabilities will separately and in combination impact the provision of incentive systems and reinforce the characteristics of the normative regulation system, whether or not it is efficient in generating social stability. Institutional development is path dependent in the long-run insofar as it feeds back on the collective learning processes and the diverse experiences that are being built in the transition from one historical moment to the next (North, 1990).

⁴ “[...] institutions not only constraint feasible strategies, but they also constitute the important players of the game and shape their perception and valuation of outcomes in the payoffs matrix” (Scharpf, 1997, p. 40)

⁵ “The present and the future are connected to the past by the continuity of a society's institutions. Today's and tomorrow's choices are shaped by the past. And the past can only be made intelligible as a story of institutional evolution” (North, 1990, p. vii).

Consequently, each society will find, among the available alternatives, its own responses to the problems it confronts in its process of historical evolution, this will have a direct impact on the definition of social skills, beliefs systems and expectations. The need of responding to critical dilemmas, many of which have to do with the very survival of the social arrangement, will shape the characteristics of the institutional framework that appears after a process of bargaining and compromising. We assume the idea that human interaction develops in the context of a particular social contract which is not to be understood as an ahistorical metaphor, but, instead, as the result of a negotiated agreement that is structured within the traditions and values and historical developments of a given group of people. Social contracts, then, are concrete phenomena that are defined by the very nature of human associations and the institutional framework that permit the society to respond to the fulfilment of common interests (Binmore, 1990).

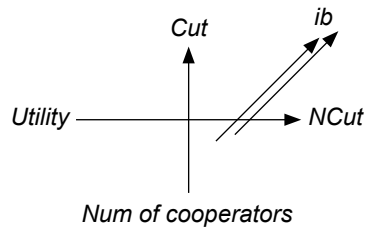
Any process that implies the alteration of the nature of public association (such as constitutional reforms, changes in the political regime, declarations of independence, decolonization processes, transitions to democracy, massive immigration, war, rebellion, emancipation) would require a renovation on the structure of human cohabitation or its transformation which can be considered as a re-definition of the term of the social agreement that will modify the rules of the game and the payoffs system. Otherwise, society could confront social turmoil and structural disorder. Such an arrangement is defined as a process of structuring and re-structuring the contents of the social contract. Those processes may or may not be successful. The social contract is, then, a bargaining equilibrium solution for social institutions that will set the parameters for determining the scope of social interactions and its efficiency levels. We are, after all, talking about the conditions that will favour, or not, the construction of social agreements that will produce political stability and will affect its permanence over time.

Searching for equilibrium

Let us consider the Stag Hunt Game argument for social cooperation under conditions of strategic interactions:⁶ in a primitive group of hunters, players have two alternatives; they can hunt a Hare or a Stag. Hunting the former is relatively easy; it can be done individually as it does not require the collaboration of any other member of the group. Hunting a stag, on the other hand, is much more complex, but

⁶ We are adopting the Stag Hunt metaphor because the structure of the game is useful to understand processes of “adoption or modifications of the Social Contract for mutual benefits” (Skyrms, 2004, p. 9).

also more rewarding for the group.⁷ The animal needs to be chased and ambushed by several individuals at the same time. A Stag is more willing to fight than a hare, and much better equipped to do so. A single person would not be able to trap such a large animal by herself. That endeavour only would be possible with the active participation of other hunters. Certainly, a Stag would cover more efficiently, and for several days, the instrumental feeding requirements of the group,⁸ however the hare would easily respond to the immediate needs of each hunter acting individually. The situation implies a contradiction between the interests of each individual and the interests of the collective that needs to be worked out. To solve this social dilemma, it would be helpful to implement an equilibrium solution based on cooperative behaviour.⁹ It is, after all, broadly accepted, by the relevant literature, that cooperative behaviour produces a higher rate of net social benefits, while, at the same time, has a potential positive impact in individual welfare. Non-cooperative behaviour is inefficient as it produces a more reduced outcome. Increments in the number of cooperators increases social utility (ut) and *vice versa*.



There is a direct relationship between increments in the number of cooperators and increases in Utility (Cut). Non-cooperative Utility distribution (NCut) is represented as producing fewer net benefits. However, we can see a gap between Cut and NCut that represents the appropriation of ut that any individual can obtain by observing a non-cooperative behaviour. Individual benefits for non-cooperative behaviour will reduce the incentive for cooperation in the long run, thus affecting the coherence of the social enterprise.

⁷ We are assuming a society of free and equal individuals, interested in satisfying their basic needs of food and shelter. The Stag will provide a higher level of calories to the entire group; however, a hunt for the hare will suffice to satisfy the individual needs of a single hunter. It is assumed, however, that the stag will have a higher return to the individual hunter even though it will signify a higher degree of effort in coordination.

⁸ A stag will respond more fully to the calories required by the members of the group. We are using the metaphor exclusively referred to the fulfilment of proteins needed by the society. It is not our interest in this case to discuss the internal "power games" that may or may not develop about the hunting process.

⁹ Certainly, the situation can be solved by the imposition of a distributive criteria by an authoritarian decision maker without considering the existence of the multiple interests that may be present in social interactions. We are interested, however, in a pluralistic autonomous solution, and that will be the centre of our argument.

J.J. Rousseau in his “Discourse of Inequality” first introduced the Stag Hunt Game (STG) as a philosophical dilemma. Brian Skyrms (2001) has noted the importance of STG for understanding the problems of social evolution. As the story goes, hunters face a predicament: they need to decide whether to join a cooperative effort or, simply, advance an individual course of action considering the potential results of each strategy. Once the players have decided to hunt the stag, one would expect they will agree on a course of action: they will follow the trail left by the animal, they will keep their posts, and they will lead the stag to the killing site. If only one of those involved in the process fails to fulfill these duties, the stag will escape leaving all of them without prey. That will produce the worst collective and individual possible outcome for those who cooperate. The Stag Hunt Game has two dynamic equilibria. One is associated with cooperative behaviour while the other with non-cooperative behaviour. In one case, the stalkers will keep their promises by cooperating with the others, thus favouring their collective interest; in the other case, the hunters will retreat from the agreement and will proceed by their own means. In each case the players will receive an outcome that results from their interdependent choices, as it is shown in the following chart.

Stag Hunt Matrix of Payments

	Stag	Hare
Stag	4/4	0/3
Hare	3/0	3/3

Individuals, then, decide whether or not to contribute to collective action. In case they do, they will receive 4ut each. In the case they do not they will receive 3ut. It is worth considering that if one of the players tries to hunt a stag on their own, or is betrayed, she will receive 0ut as the animal will escape while her counterpart will receive 3ut as a reward for hunting a hare. We assume the hunter calculated the opportunity cost of not hunting a hare against the benefits of cooperative hunting. Betrayal, in the long run, will discourage cooperative behaviour and most likely produce a less efficient equilibrium. Many social games, as the one represented here, have more than one available equilibrium. N-Players games face important difficulties in reaching efficient agreements, as they require the evolution of conventions for the coordination of behaviour.¹⁰ Each society will develop a grammar that will allow its members to: communicate with each other; build agreements; understand abstract signals; negotiate; and learn (Bicchieri,

¹⁰ “Convention is a pattern of behaviour that is customarily expected, and self-enforcing. Everyone conforms; everyone expects others to conform given that everyone else conforms” (Peyton, 1993, p. 5).

2006; Skyrms, 2010). Social games are cooperative in the sense that they entail the possibility of making binding agreements (Harsanyi, 1986). However, communication is only possible if we recognize the content of the network of messages and information built into human interaction and exchange (Kahneman, 2011; Li, 2009). This process contributes to creating patterns of normative behaviour that need to be recognizable for other members of the society. Conventions will help us to define expectancies about the behaviour of others, reduce uncertainty and make predictions about the scope of the game (Lewis, 1969),¹¹ particularly in the case of large worlds (Binmore, 2011).

We use N-Players Games as an analytical model for complex strategic interactions where the interests of players are not *perfectly coincident* (vid Schelling, 1960). We assume variability in their values and preference systems. Each society must define the terms of the game of social interactions, both public and private, as it implies the distribution of costs and benefits and the definition of incentive mechanisms. Consequently, they must determine how to organize to: attend the dynamics of human interaction; make productive processes viable; maintain social order; guarantee economic exchange, educate its youth; make health care available to those who are in need; integrate those who are excluded; and, establish a level of decency in the way state institutions do their job and guarantee accessibility to the “basic goods” that are socially required (Rawls, 1999).¹² In other words, each society needs to solve the Equilibrium Selection Problem (Binmore, 2007). The collective dilemma of structuring a social order can be solved by means of cooperative or non-cooperative solutions, each one of which produces its own particular stability. As we saw in the Stag Hunt Game, there are two possible equilibria. One in the upper left square and the other one in the bottom right. The former is a representation of cooperative behaviour while the other denotes non-cooperation. Games with greater complexities may have more than one cooperative or non-cooperative equilibria. Let us consider the following case: A and B need to meet to solve a situation. Due to time schedules and distances there are three possible places they can meet with equivalent positive utility distribution, while there are several others which they will not choose because they are inconvenient, at least for one of them as they result in an inefficient outcome at least for one of them. Please note that, at this point, we are considering a cooperative game, so that they can communicate and make binding agreements.¹³

¹¹ “Norms serve to guide individual’s behaviour, but also to allow an individual to anticipate other’s behaviour... a social norm is, in a way, a cluster of expectations” (Bicchieri, 1990, p. 840).

¹² A minimal level of decency will also be necessary in terms of private or public interaction among players. They should be able to understand each other but also, they need to be able to respect and to feel empathy for others, as they are qualifying as equals in terms of their human condition.

¹³ For a discussion about the problem of coordination and convention, see Lewis (1969/2007).

		A		
		a 2/2	b 3/0.5	c 4/0
B	d 1/0	e 1.5/1.5	f 0/1	
	g 0.5/2	h -1/-2	i 1/1	

As A and B are rational players, we will suppose they will meet wherever they both receive a positive outcome and will reject all other possibilities. Consequently, their meeting will be held in places labelled a, e or i, but nowhere else. Despite the fact that $a > e > i$, any one of these three possibilities will be acceptable as there is a non-objectable equilibrium point. Nonetheless, it is worth nothing that they still must negotiate the meeting place among those available. One will suppose that given the utility distribution they will meet at ‘a’; that, however, will not, necessarily, be always the case. There may be cultural, ideological, or appraisal conditions that may lead them to a less efficient outcome, in this sense, rationality is bounded (Selten, 2002). In the Stag Hunt Game, players will accept the less efficient non-cooperative solution, if they consider that the probability for defection is high. It is worth noting that a non-deer eating community will not hunt the animal, for instance, because of a ban against deer hunting for religious reasons, or a community that historically has only caught hares, or, for that matter, a group of vegetarians. In these cases, it is much more difficult to build a cooperative solution associated to stag hunting even if it implies a higher distribution of utility. In favourable conditions inefficient solutions stay (Bicchieri, 2017), players will optimize expected utility in the context of a stable solution despite its quality. We may find, for instance, historical situations where certain social groups have ended up supporting bad governments or violent dictatorships.¹⁴ People who learn how to survive in adverse situations can take advantage of particularly inefficient institutional designs, or simply accept the “way things work”; as they get used to it they will lose their capacity to evaluate their own preferences, or interests as they may become suppressed or attenuated due to particular circumstances, such as authoritarianism or paternalism, which will reduce their autonomic capacity to decide by themselves.¹⁵

Society moves towards the definition of an equilibrium solution even if it is not necessarily the most efficient one. In fact, it is crucial to ask for the reasons that explain the permanence of institutions that clearly work against the maximization of the social utility function. Strong arguments have been made in favour of a natural

¹⁴ There is an interesting discussion about these kinds of selections in Bicchieri, 2017, particularly in chapter 1.
¹⁵ “Real people seldom distill abstract rationality principles from their experience. They simply learn what works in certain stereotyped situations” (Binmore, Samuelson. 1994, p. 51).

evolution of the social contract (Axelrod, 1984). There is, however, strong empirical evidence that Axelrod's *shadow of the future* is a necessary but insufficient condition for the evolution of cooperative behaviour, or for its survival in a repetitive game (Binmore, 1989; Bicchieri, 2006). We need to consider, after all, that equilibrium points may fluctuate as society changes. Cultural and/or historical biases as well as ideological conditions will play against the possibility of constructing win/win agreements if institutional conditions do not favour a fair distribution of utility, allow for cultural changes, and facilitate the evolution of trust. Cooperative behaviour cannot be founded on long-term coexistence, as some literature posits. People can live together in horrible conditions without the willingness to change the term of their interactions or without even perceiving that they are in disadvantageous situations. After all, they can learn how to survive, and, under particular circumstances, surviving may be considered as the only available option. *The shadow of the future* may work in an impartial *well-ordered society* based on justice, but it is less likely to do so if conditions are adverse, particularly when the winners and the losers of social interactions are always the same, when results are unfair, or are unfairly distributed. As we face terrorism, immigration, extreme poverty, and inequality in the distribution of income, among other things, we need to start thinking seriously about the perverse effects of poor institutional design in the construction of social order, particularly when public distress, low levels of participation, public protests, dissatisfaction with the 'state of affairs' and lack of support for the political system make equilibrium tremble and push towards design and distributional changes in the institutional framework.

Institutions and beyond

According to North (1994), institutions are tangible and permanent sets of rules in which evolution can be observed.¹⁶ They function as formal and informal structural parameters for social interactions, providing coherence and regulation to social interactions, and thus reducing uncertainty.¹⁷ Institutions allow us

¹⁶ "Institutions reduce uncertainty by providing a structure to everyday life. They are a guide to human interaction, so that when we wish to greet friends on the street, drive an automobile, buy oranges, borrow money, form a business, bury our dead, whatever, we know [...] how to perform these tasks [...] In the jargon of the economist, institutions define and limit the set of choices of individuals" (North, 1994, p. 53).

¹⁷ We understand Institutions as the rules of the game. They determine the possibilities available to players and the range of strategies that actors advance to achieve their objectives. They are the norms that determine the functioning structure of society, as such they "prescribe or proscribe behavior; they entail obligations and are supported by normative expectations" (Bicchieri, 2006, p. 38).

to foresee game dynamics as they define the available options for social interactions and help us to understand the historical evolution of a given society. Institutions are determinant in defining the efficiency levels of human associations. They favor, or not, the evolution of cooperation by configuring incentives systems and enforcement mechanisms. Institutions are the rules of the game;¹⁸ they structure and support the social order, as long as they help us to determine patterns of behaviour and distributive criteria. In other words, institutional frameworks are the mechanisms that shape human interactions, standardize social and economic exchange and define the boundaries between acceptable and unacceptable behaviour. Indeed, structural parameters are determinant in establishing the solutions that each society finds to the problems of order and equilibria according to its social and economic endowments, its development path, and society's perceptions about the challenges they face at any particular moment of its historical process and about the path that will lead them into a desirable future.

The institutional framework develops as a solution to the equilibrium problem of a given society. It does not tell us, however, anything about its efficiency and/or long-term stability. Certainly, social arrangements intend to properly solve social dilemmas, but not always can they guarantee the provision of order, justice, or primary goods.¹⁹ They, nevertheless, determine the characteristics of the game, its structure, and its dimension.²⁰ Institutions may lack the capability to support cooperative behaviour if the incentives mechanisms built into them send wrong messages; are incapable of applying sanctions to defectors; produce obvious unfair results; impose high social costs; and/or, apply a discriminating distribution of costs and benefits among the members of the society. They may also fail to build cooperative solutions if they do not recognize the cultural particularities of the population, their values, and their history. Unfair equilibria tend to be unstable. In such a case, order could be maintained by the use of force and/or by in-

¹⁸ “The grounds for looking at institutions as rules rest on an assumption that many observed patterns of interaction are based on a common understanding that actions inconsistent with those that are proscribed or required are likely to be sanctioned or rendered *ineffective* if actors with the authority to impose punishment are informed about them. To understand regularized patterns of interaction affected by rules, one needs to examine the actions and outcomes that rules allow, require, or forbid and the mechanisms that exist to enforce those rules” (Ostrom, 1995, p. 583).

¹⁹ “The mere presence of a social norm does not justify inferring that it is there to accomplish some social function. Besides, it does not account for the fact that many social norms are inefficient, as in the case of discriminatory norms against women and blacks or are so rigid as to prevent the fine-tuning that would be necessary to successfully accommodate new cases” (Bicchieri, 1990).

²⁰ “In short, the games that are in fact being played in policy processes are to a large extent defined by institutions” (Scharpf, 1997, p. 40).

stalling mechanisms for social control such as the limitation of civil liberties, the reduction of plurality and the lack of guarantees classically used by dictatorships; or, by manipulating the popular masses through ideology, manipulation, promises and direct subsidies, as left and right populist regimes tend to do.

The consolidation of democracy requires the development of the condition necessary for the development of cooperative behaviour, after all democracy is much more than a simple vote mechanism, it is in fact a “way of life” based on plurality, respect and social collaboration. At first glance one could say that it would be possible simply by changing the distribution of utilities among the players, however it will be necessary to consider not only incentives systems, but also values, attitudes and the disposition of the population towards the moral requirements that are required for democratic coexistence. In the Stag Hunt example, it would suppose the introduction of incentive that will induce others to maintain the cooperative behaviour, so that they will keep hunting the stag no matter the availability of hares. This change in the structure of incentive requires a fair distribution of the prey, recognition of individual and collective efforts and a positive perception about the value that cooperative behaviour has for society so that everyone perceives not only that her participation in STG is required but also that it is valuable and necessary. They will then participate in the hunting process out of the conviction that they ought to do it. These decisions will change, of course, the very nature of the game and will imply a modification of its structure.

Stag Hunt Evolved Matrix of Payments

	Stag	Hare
Stag	5/5	0/-1
Hare	-1/0	-1/-1

In this case, a sanction has been introduced by changing the distribution of utility in the lower right cell, but also there has been an increase in the utility associated with the cooperative solution, indeed it has been transformed into a dominant solution, which implies that hunters will, in all cases, contribute their part in hunting the stag despite the presence of hares,²¹ all of this under the general perception that this solution is a fair one. As we have noted, these changes will impose increased costs on defection and will induce actors to cooperate, or otherwise to pay the cost of defection. By changing the incentive system, we have built only one equilibrium solution into the game. Theoretically, it will increase the utility function of hunters. As noted before, changes in the

²¹ We assume that in all cases the community has the capacity to impose more than a proportional sanction on those who deviate from cooperation.

distribution of utility alone does not appear to guarantee the stability of the social arrangement, our actions, after all, are not dependent on material revenues alone, in fact in many cases non-tangible incentives, such as recognition or the satisfaction of having fulfilled our duty, play a very important part in the definition of our behaviour. Material incentive alone or even threat of punishment, will induce changes of behaviour by external forces, people may feel obliged to comply so that there could be an increase in the coherence of collective action. Notwithstanding, changes in payment distribution will be irrelevant unless they become encompassed with variations in the distribution of expectations and beliefs. Those incentives will favour the development of a compromise with the moral duty to cooperate with others in the achievement of a common goal. In this case a transition from a type of rationality based on individual rationality towards a cooperative rationality will be in place. The subject will restrict, due to her own conviction and sense of duty, her individual maximization.

In general, the development of a fair equilibrium solution is the result of large processes of trial and error that can be mediated by social unrest, protests, political manoeuvres and negotiation. Cooperative solutions require the construction of agreements and the aggregation of interests among the players; this will allow them to determine, in a process of social aggregation, the content of fairness and its impact on their well-being. This process is equivalent to building a bargaining equilibrium solution that considers the interests, values, and characteristics of those who take part in subsequent interactions. That is, the interested parties need to agree on the contractual structure of their long-term interactions, not only based on their bargaining strengths and interests, but also on their shared principles, values and learning processes. This will be possible when the agreement results from a fair bargaining process and a common criterion of justice. These elements will be required as justificatory mechanisms with a direct impact on the validation of moral standards.

Indeed, the stability of the agreement will depend on its effects on the associates and on the legitimacy of its origins. Results perceived as fair will produce a relatively permanent equilibrium while biased or unfair ones will put pressure on institutional change. Successful changes have resulted, for instance, in the case of the recognition of civil rights movements in the USA during the second half of the last century; on women's liberation and rights equivalence; or, with the contemporary reception of migrants' in Europe. These processes required negotiations that changed the distribution of utility among the concerned parties as the opportunity costs of changing the social arrangement became smaller than the costs of maintaining expensive resistance to change or system instability. They, also, demand the development of a positive valuation of the transformative dynamics that entails a moral betterment in society as well as an increase

in equality, social inclusion, recognition, autonomy and self-respect for those favoured by the transformation of social relations. These advancements come in many forms, for example: when most of the population understand that slavery is not only illegal, but is also against the moral norms of humanity as it is universally accepted that no human being has the right to own another one; rape or any other form of violence becomes unjustifiable; and, child labour, gender inequality, and racial segregation start being considered as morally wrong and, consequently, unacceptable.

These are crucial developments in the construction of social order, as morality becomes a normative mechanism for underpinning the transformation of payment systems. Change will be shaped not only when players realize that they are unwilling to pay the cost of unstable equilibrium, but also when it becomes clear that the results of social interactions are frankly unfair, or that certain groups are treated unfairly due to their gender, race, ideologies, or world views. Preferences, and belief systems, are not necessarily fixed, and they do not depend on power distribution. Indeed, they can change in the presence of positive incentives, including cultural variations. The confluence of incentives related to changes in the distribution of utility as well as changes in the environment may have an impact on the way people perceive and value their reality. After all, humans adapt to environmental changes by adopting strategies that will help them to cope with the changes more efficiently (North, 1994), but also there is a tendency to value the consequences of social choice in terms of its extended consequences over the population. Group membership guarantees a minimum of considerations towards our needs and preferences. But above all, nobody will rationally bet for a social arrangement that produces a less than stable equilibrium, even though, as we have said before they can adapt to live in less than favourable conditions.

Morality evolves as a strategic mechanism for the consolidation of social order. In fact, it reinforces the functionality of the institutional framework. We are not talking, only, about sanctions and incentives represented by utility distribution, instead, we are referring to the validation and acceptance that society grants to the normative arrangement and its consequences. Society develops as a result of a learning process that will produce evolutive advantage based on the most efficient negotiated solution to social dilemmas. This solution would be associated with the reduction of uncertainty, the limitation of conflict, and the emergence of cooperative behaviour and its permanence in the long run. Players will end up understanding that to act according to a moral conception, widely known and accepted by the members of that society,²² is not just a pragmatic exercise, but furthermore a response to the question of what ought to be done in

²² We understand that both moral and legal standards need to be common knowledge.

terms of their interactions with other members of the society and the expected result of such a process based on reciprocity and fair distribution.

Cooperative equilibrium does not respond to altruism nor naivety, but, instead, to a process of awareness about the advantages associated with the reduction of the transaction costs bred by the reduction of conflict and the distribution of cooperative surpluses. As an alternative for maximizing their individual utility in the strategic setting, rational players would produce an agreement that will enable a stable maximization of social utility under suitable conditions.²³ We must understand the individual participation in the social contract as an act of will. There is, after all, a difference between coordination of behaviour and social cooperation. The former can result from a compulsive obligation hierarchically imposed over the members of the society despite their values, beliefs, or interests; punishment will be used to correct deviations, but even more autonomy will be suppressed or, at least, attenuated, as individuals become dependent on the will of others: dictators, fathers, husbands, ayatollahs, owners, masters.

Cooperation, on the other hand, requires an act of the will. Cooperative behaviour, then, answers not only to self-interest, or external pressures, but also to a positive individual valuation about the moral worth of a given solution. It is interesting to note that in the case of coordination people have no alternative but to harmonize, as required, their behaviour with other members of society. They are, simply, compelled to do so, without considering their own preferences. In the case of cooperation, individuals can choose how to act. They will cooperate out of the conviction that, under the adequate conditions, it would be proper to do so. In the Stag Hunt Game, we could obligate the hunters to keep their posts and restraint from hunting a hare by imposing sanctions on non-cooperative behaviour, that will require permanent surveillance and monitoring. Authoritarian solution, however, seems to be too expensive and inefficient as it implies permanent supervision of the behaviour of others and the deployment of a long-range sanction system difficult to manage and to finance. In such a case, society will face the risks of corruption, discretionary behaviour, and loss of legitimacy and control.

The cooperative definition of a normative order would result from a process of negotiation and permanent adjustments that deals with the problems of previous power distribution, resources, and asymmetries. It recognizes fair distribution, equality, and justice as required foundations. Once in place, agreements must be adjusted as contextual conditions change, including changes in social preferences and progression or regression in the moral standing of the society.²⁴ It is important to reinforce the idea that preferences are not fixed; they can

²³ We are assuming here that under the proper conditions rational players are moral players.

²⁴ Evolution is not necessarily unidirectional.

change when the appropriate incentive exists. For instance, when people worry about fairness, they will prefer a just solution to an unfair one (Bicchieri, 2006). But, the perception of rightness has to do with the evolution of moral values as they become right or not for social evolution. Race segregation and gender inequality, for instance, were considered acceptable until not long ago. The same can be said about massive environmental destruction or school harassment (bullying). As the game develops, a moral transformation in social values takes place. People start to value positively, or negatively, the transition to a new institutional structure and payment system in complex feedback processes that sets the paths for institutional change or permanence. It seems to be clear that in complex social settings conformity to rules does not depend only on sanctions, but also on the disposition of individuals to accept and validate the institutional boundaries prescribed by social norms. In other words, if institutions are going to facilitate the development of cooperative behaviour it is necessary for cooperation to be previously considered as a valuable form of behaviour.

Conclusions

As we have seen, consideration about the value of cooperative behaviour goes beyond both economic cooperative surplus, positive incentive and/or sanctions. Cooperation takes root in the working structure of social agreements based on fairness and convictions. In the Stag Hunt example, it appears to be clear that sanctions would not be enough to guarantee the confluence of cooperators, unless they felt an individual disposition to act according to a common goal once it has been traced. Otherwise the *temptation* of defecting from the cooperative effort will be constantly present. Hunters will defect from Stag Hunt whenever they have an opportunity to do so, unless they are convinced that they should not and there are incentives in place to reinforce cooperative behaviour. It is not a naïve conviction but, instead, one based on rational evaluation and appropriation of moral standards. The hunters will consider cooperation as an equilibrium solution under the proper conditions and equity in the distribution of costs and benefits. After all, conflict of interests needs to be solved to guarantee the coherence of social interactions and exchange; every solution will need to be based not only on the capacity to impose sanctions on defectors, but also on the positive evaluation about the characteristic of the institutional framework and its consequences, including in this evaluation the moral foundations that support the agreement. This results from a process of learning by doing. Trial and error will make clear the necessity of bargaining in favour of stable social settings. As we have argued, social dilemmas cannot be solved merely by imposing costs on non-cooperators

or by implementing new legislation. Certainly, sanctions will reduce free-riding and minimize and facilitate the coordination of behaviour, however they cannot assure players' willingness to cooperate in fulfilling collective objectives.

Cooperation, indeed, responds to one's disposition to accompany others in achieving common goals. That, by the way, implies the restriction of individual maximization in order to build agreements with others considered as valid members of the society and our moral equivalents as members of the human race (people with whom we will speak freely and in equal conditions, people we respect and with whom we will sign a contract if we have the opportunity to do so). Cooperation, then, is a public action that implies an act of will that is manifested by intentional and visible activities. Such a disposition is justified by the expectation of individual gains in the context of a stable equilibrium that allows people to foresee long-term benefits. Cooperation requires moving away from individual maximization to favour collective well-being in the concrete and rough world in which we live and beyond an idealistic solution. Just as Binmore and Samuelson refer: we must, indeed, learn to build bargaining solutions in the imperfect world we live in. That would require the conviction that it is the right thing to do. This process can be interpreted as the evolution of a moral validation for social norms, and both collective and individual behaviour.

Bibliography

- Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.
- Barragan, J. (1998). "Ethical Issues in Public Decision Making." *Rechtstheorie*, 29, pp. 23–45.
- Bicchieri, Ch. (1990). "Norms of Cooperation." *Ethics*, 100(4), pp. 838–861.
- Bicchieri, Ch. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.
- Bicchieri, Ch. (2017). *Norms in the Wild: How to Diagnose, Measure and Change Social Norms*. Oxford: Oxford University Press.
- Binmore, K. (1989). "Social Contract I: Harsanyi and Rawls." *The Economic Journal*, 99(395), pp. 84–102.
- Binmore, K. (2011). *Rational Decisions*. Princeton: Princeton University Press.
- Braithwaite, R.B. (1954). *Theory of Games as a Tool for the Moral Philosopher*. Cambridge: Cambridge University Press.
- Buchanan, A., Powell, R. (2018). *The Evolution of Moral Progress: A Biocultural Theory*. Oxford: Oxford University Press.
- Cohen, G.A. (2003). "Facts and Principles." *Philosophy and Public Affairs*, 31(3), pp. 211–245.
- Forst, R. (2008). "Toleration and Trust: Comments on Steven Smith." *American Society for Political and Legal Philosophy*, 48, pp. 281–292.

- Gauthier, D. (1986). *Moral by Agreement*. Oxford: Oxford University Press.
- Goodin, R., Tully, Ch. (2006). *The Oxford Handbook of Contextual Political Analysis*. Oxford: Oxford University Press.
- Güth, W., Ockenfels, A. (2005). "The Coevolution of Morality and Legal Institutions: An Indirect Evolutionary Approach". *Journal of Institutional Economics*, 1(2), pp. 155–175.
- Harsanyi, J. (1986). "Individual Utility and Utilitarian Ethics". In: A. Diekmann, P. Mitter (eds.). *Paradoxical Effects of Social Behavior: Essays in Honor of Anatol Rapoport*. Heidelberg–Wien: Physica-Verlag.
- Kahneman, D. (2011). *Thinking Fast and Slow*. London: Penguin Books.
- Kühnelt, J. (2010). "State Legitimacy and Social Order". In: J. Kühnelt (ed.). *Political Legitimation without Morality?* Dordrecht: Springer, pp. 15–33.
- Lewis, D. (1969/2007). *Convention: A Philosophical Study*. Oxford: Blackwell Publishers.
- Li, M. (2009). "Collection in Bargaining". *International Journal of Game Theory*, 38, pp. 481–498.
- North, D. (1990). *Institutions, Institutional Change and Economic Performance: Political Economy of Institutions and Decisions*. Cambridge: Cambridge University Press.
- North, D. (1994). "Economic Performance through Time". *The American Economic Review*, 84(3), pp. 359–368.
- Ostrom, E. (1998). "A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997". *American Political Science Review*, 92(1), pp. 1–22.
- Peyton, Y. (1993). "The Evolution of Convention". *Econometrica*, 61(1), pp. 57–84.
- Rawls, J. (1977). "The Basic Structure as a Subject". *American Philosophical Quarterly*, 14(2), pp. 159–165.
- Rawls, J. (2004). *Teoría de la Justicia*. México: Fondo de Cultura Económica.
- Scharpf, F. (1997). *Games Real Actors Play: Actor-Centered Institutionalism in Policy Research*. Oxford: Westview Press.
- Schelling, T. (1956). "An Essay on Bargaining". *The American Economic Review*, 46(3), pp. 281–306.
- Schelling, T. (1964). *La Estrategia del Conflicto*. Madrid: Editorial Tecnos.
- Selten, R. (2002). "What Is Bounded Rationality". In: G. Gigerenzer, R. Selten (eds.). *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.
- Sen, A. (1999). "The Possibility of Social Choice". *The American Political Science Review*, 89(3), pp. 349–378.
- Sen, A. (2000). *Development as Freedom*. New York: Anchor Books.
- Skyrms, B. (1996). *The Evolution of the Social Contract*. Cambridge: Cambridge University Press.
- Skyrms, B. (2001). "The Stag Hunt". *Proceedings and Addresses of the American Philosophical Association*, 75(2), pp. 31–41.

- Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure*. Cambridge: Cambridge University Press.
- Skyrms, B. (2010). *Signals, Information, Learning and Information*. Oxford: Oxford University Press.
- Wagner, E. (2012). "Evolving to Divide the Fruits of Cooperation". *Philosophy of Science*, 79(1), pp. 81–94.