The Institutional Dimension of Salience: Common Understanding and Embeddedness

Cyril Hédoin
OMI, University of Reims Champagne-Ardenne

3^{nd} version - Do not quote without permission

September 9, 2011

Since Thomas Schelling's pioneering work, salience and focal points appear to be useful concepts to explain how coordination is possible. However, game theory has struggled for forty years in trying to give to salience a formal expression. Recent works have succeeded in integrating salience in formal models. However, these works do not explain the origins of salience. This paper argues that salience has an institutional dimension. It builds on and expands David Lewis' theory of common knowledge by interpreting salience as a public phenomenon where individuals know that they share background information and practical reasoning. More specifically, we show that salience is a public and social phenomenon which occurs when individuals have a common understanding of a given situation. This common understanding results from the fact that individual action is embedded in social relationships organized by institutions.
JEL classification: C72, C73, D83.
Keywords: Salience, common knowledge, institutions, we-reasoning, common understanding, embeddedness.

In what is arguably a seminal work in strategic thinking, Thomas Schelling [Schelling(1960)] emphasizes humans' ability to coordinating their actions in settings where no obvious device to choose appears to the observer. Schelling famously coined the terms of "salience" and "focal points" to explain this proclivity. Economists have struggled with the idea of salience since Schelling's pioneering work. Two problems with the salience concept help to explain this fact: first, at the theoretical and formal level, salience is difficult to insert in game-theoretic models and consequently is perceived as an *ad hoc* assumption not contributing to enhancing the explanatory power of the theory. Second, at the cognitive level, the working of salience remains somewhat mysterious. Because of these two difficulties, economists have generally tried to escape from the concept of salience or to derive it from more fundamental mechanisms.

The purpose of this paper is to develop an explanation of salience by demonstrating its institutional dimension. We argue that any explanation of salience must take into account its public and collective aspects. The main point is to show that the empirical significance of salience in economic interactions cannot be captured by a uniquely formal game-theoretic analysis but calls for a reflection on the institutional mechanisms underlying this phenomenon. Building on David Lewis' ([Lewis(1969)]) theory of common knowledge, we point out the necessity to explain how individuals may have confidence in the fact that they share the same background information and similar practical reasoning. We suggest that the sociological concept of embeddedness ([Granovetter(1985)] can help to achieve this end. We show that the institutional embeddedness of individual action helps to foster a common understanding of the situation among the protagonists. This common understanding allows a form of collective reasoning through a fictitious stage where individuals 'agree' to make some features of the situation salient, permitting the common knowledge of expectations. This tacit agreement defines *public* salience.

The paper is organized in eight parts. The first part discusses some of the puzzles in coordination games and mixed-motives games which economists have faced for a long time now and explain why salience might be an answer to them. The second part introduces David Lewis' theory of common knowledge and explains the role Lewis gives to salience. We argue that in Lewis' theory, the property of salience is to generate a common knowledge of beliefs and expectations among participants in an interaction. The third part defines salience as a special kind of correlated equilibrium. A short digression in the fourth part notes that this makes salience formally equivalent to assume that players have common priors. The fifth part argues that salience (and hence common priors) has an institutional dimension that makes necessary to go beyond the individual level. We expand on this last point in the last three parts: we characterize salience as a public phenomenon and we establish that it presupposes common understanding of the situation among the members of a population. Finally, using the concept of embeddedness, we demonstrate how institutions can help to foster this common understanding. A short conclusion suggests that the necessity to integrate public salience in game-theoretic models is one of the many signals indicating that methodological individualism is a logical dead end.

1 Indeterminacy and Salience in Game Theory

Economists and game theorists continue to struggle with two kinds of problem related to solution concepts in game theory: multiplicity and indeterminacy ([Hargreaves-Heap and Varoufakis(2004)], [Sugden(1991)]). The problem of multiplicity results from the fact that many games (except those with a dominant strategy) have several (pure and mixed) Nash equilibria.

Multiplicity poses a serious threat to game theory conceived as a predictive and a normative tool. The Nash equilibrium refinement research program ([Harsanyi and Selten(1988)]) has dealt with this problem but has remained largely unsuccessful. Incidentally, this failure provided one of the main motivations for the development of new forms of game theory, such as evolutionary game theory ([Samuelson(2002)], [Sugden(2001)], [Young(1998)]). Indeterminacy is a logical consequence of the multiplicity problem: since there are often more than one Nash equilibrium in a game, the outcome of the game becomes impossible to determinate on *a priori* grounds. However, the problem of indeterminacy expands beyond this limit: the theorist is not the sole person to face indeterminacy; in fact, the tacit assumption of a symmetry between the way the modeler describes the game and the way the players frame the game they are playing implies that players also face indeterminacy. This second limit is severe because then it is simply impossible for the theorist to explain how and why players behave as they do in a specific game. In fact, when there is indeterminacy, players themselves must be unable to choose. Indeterminacy goes even further once we realize that the traditional assumption of common knowledge of rationality among players does not entail that one of the Nash equilibrium will be played. For the latter to be the case, players must also have common priors on the way the game will be played by each one of them. This assumption, first formalized by Robert Aumann ([Aumann(1987)]), is at best controversial ([Hargreaves-Heap and Varoufakis(2004)], [Sugden(1991)]).

The difficulties resulting from the problems of multiplicity and indeterminacy can be illustrated through the "Hi-Lo game". Despite its simplicity, it reveals the failure of conventional game theory to be both an explanatory and a normative tool:

	A	B
A	2 ; 2	0 ; 0
B	0 ; 0	1 ; 1

That both players should play "A" seems obvious given the fact that the [A;A] equilibrium is payoff-dominant. Indeed, experiments show that a large majority of individuals are able to coordinate on this particular equilibrium ([Bacharach(2006)]). However, if we make the traditional assumption that players are instrumentally rational and that this fact is common knowledge among the players, we are unable to explain this simple empirical observation. It is clear that the players cannot avoid engaging in an infinite regress reasoning: I should play A if I think that you will play A; but you will play A if you think that I will play A; so, I will play A if I think that you think that I will play A, and so on. This chain of reasoning never ends and it can be developed also for strategy B. Nothing in the mathematical description of the game and in the way rational players reason can help them to choose between A and B.[1] Therefore, we are left with a

[1]Note that to allow mixed-strategies is not a solution. First, the traditional understanding of mixed-strategies has been criticized as unrealistic. Second, if we follow Aumann ([Aumann(1987)]) in interpreting mixed-strategies as beliefs on the way the other player will play, it is problematic because if I believe

theoretical puzzle and a failure to explain an empirical evidence.

In the third chapter of *The Strategy of Conflict*, Thomas Schelling presents the results of several informal experiments building on similar game structures. Experiments include both (pure) coordination games where the interests of players are perfectly aligned, and mixed-motive games where interests partially conflict while an element of coordination is preserved. Schelling's aim was to show that in games where the mathematical structure seems to leave players unable to make any choice but a random one, individuals demonstrate an ability to coordinate on some specific outcomes. As he puts it (p. 57):

> People *can* often concert their intentions or expectations with others if each knows that the other is trying to do the same. Most situations - perhaps every situation for people who are practiced at this kind of game - provide some clue for coordinating behavior, some focal point for each person's expectation of what the other expects him to expect to be expected to do... A prime characteristic of most of these "solutions" to the problems, that is, of the clues or coordinators or focal points, is some kind of prominence or conspicuousness.

According to Schelling, people are able to "solve" coordination and mixed-motive games because they rely on *focal points*. Focal points are defined by their "conspicuousness", their prominence, or their salience. Prominence is linked to some kind of uniqueness characterizing the solution. This uniqueness itself depends of the way individuals see or frame the interaction, as Schelling makes it clear by arguing that prominence is more a matter of "imagination" than of "logic": "in the final analysis we are dealing with imagination as much as with logic; and the logic itself is of a fairly casuistic kind" (p. 58). Schelling insightfully argues that communication alone does not provide the salience needed to solve problems of coordination, even in pure coordination games. This is so because coordination is principally a matter of *expectations* rather than of actions. Communication can help to make the discovery of focal points easier but is not a substitute to them. What is decisive is that expectations converge toward a definite point to form mutual expectations.

Schelling's most famous informal experiment indicates that in a situation where one would have to meet somebody in New York City at a place and at an hour to be determine without prior communication, a majority of individuals would have chosen to go at Grand Central Station at 12 noon. He cites also experiments in mixed-motive games showing the importance of prominence in determining the outcome. For example, he used the "heads or tails" game with the following payoff structure:

---

that you will play a mixed-strategy, then by definition at the equilibrium I am indifferent between any pure and mixed-strategies. Moreover, we have to determine from where these beliefs come from. This is again the problem of the common prior assumption noted above.

	Heads	Tails
Heads	3 ; 2	0 ; 0
Tails	0 ; 0	2 ; 3

Schelling indicates that 16 out of 22 row players and 15 out of 22 column players chose "heads". Interestingly, despite the fact that this equilibrium advantages row players, a large majority of column players were led to choose "heads" because they expected that row players would hold the same expectations. Therefore, it seems that salience can help to overcome conflicting interests, at least when the conflict is not too great. [Mehta, Starmer, and Sugden(1994b)] and [Mehta, Starmer, and Sugden(1994a)] have reproduced some of Schelling's experiments (though for coordination games only) in a more formal way and found similar results. Given the fact that most of the games individuals play in the economic world are coordination games and mixed-motive games, the empirical significance of salience cannot be disputed. It urges economists to provide a satisfactorily account of salience and focal points.

Schelling provides an explanatory account of salience and focal points in the fourth chapter of *The Strategy of Conflict*. Schelling's explanation mainly relies on the idea of *framing*. As we have already noted, Schelling makes the perceived uniqueness of an outcome or of an action one of the main feature of salience. More generally, Schelling points out the fact that to understand salience one needs to distinguish the way the analyst describes an interaction through the mathematical properties of a game and the way participants perceive this very same interaction:

> We must avoid assuming that everything the analyst can perceive is perceived by the participants in a game, or that whatever exerts power of suggestion on the analyst does so on the participant in a game. (p. 113)

The mathematical properties of the game are only one of many aspects that contribute to define an interaction from the point of view of people participating in it:

> It is that the mathematical properties of a game, like the aesthetic properties, the historical properties, the legal and moral properties, the cultural properties, and all the other suggestive and connotative details, can serve to focus the expectations of certain participants on certain solutions. (p. 113)

To explain salience, the theorist must take into account any feature of an interaction that can generate stabilized convergent expectations on the way everyone will act. Schelling underlines two separate aspects that contribute to this understanding: labeling and pattern recognition. Labeling consists in identifying in specific ways rows and columns strategies as well as players such as they are meaningfully distinguishable. As Schelling himself notes, since the pioneering work of Luce and Raiffa on game theory ([Luce and Raiffa(1957)]), it is traditionally assumed that labeling of strategies and players should not matter; they are neutral descriptions used by the modeler for expository

purpose. But Schelling notes, "[i]t is precisely because strategies are "labeled" in some sense - that is, have symbolic or connotative characteristics that transcend the mathematical structure of the game - that players can rise above sheer chance and "win" these [coordination] games" (p. 96). Because individuals can label strategies and participants in a variety of ways, what look like identical games on mathematical grounds can in fact be very different situations, leading to significantly different outcomes. Therefore, Schelling points to the necessity of understanding the labeling process to study games.

Pattern recognition is a different but related issue. Many real games are repeated one. They may involve the same participants or not. Anytime an individual will be involved in an interaction that bears some resemblance with a preceding one he has been part of, he will rely on significant features that may help him to predict how things will fare. In a bargaining situation, past behavior of partners may conduct some signal and information on the way they will behave in the future. Any participant in a repeated interaction searches for recurring patterns that can help him to act optimally the next time. However, patterns are not objective features but rather subjective perceptions. They are not captured by the mathematical structure of a game but again depends both on the characteristics of players and of the interactions: "Presumably, [individuals] find their patterns in such things as natural boundaries, familiar political groupings, the economic characteristics of state that might enter their value systems, Gestalt psychology, and any clichés or traditions that they can work out themselves in the process of play" (p. 104).

2 Salience and Lewis' Theory of Common Knowledge

Some game theorists have recently tried to integrate Schelling's idea in a formal framework by explicitly modeling salience as a result of framing. For example, [Sugden(1995)] develops a model where focal points are the result of the 'labeling functions' ascribe to the players. The basic idea is that any individual in a strategic interaction is necessarily "choosing under a description": each strategy available is given a label making able players to distinguish it. Similarly, Michael Bacharach' s variable frame theory ([Bacharach(1993)], [Bacharach(2006)]) assumes that individuals play games using a particular frame assigning one or several descriptions to each strategy. A salient outcome is an outcome which has a particular uniqueness under a particular frame. Both Sugden's and Bacharach's theories provide a formal framework to explain how salience can help individuals to choose. However, since labels and frames are exogenous in these frameworks, we do not have an explanation of salience. The same applies to the way the philosopher David Lewis ([Lewis(1969)] uses the salience concept. However, Lewis develops a theory of common knowledge which helps us to gain a better understanding of the epistemic foundations of salience. In turn, these epistemic foundations will allow us to go beyond the idea of salience as a mere individual phenomenon.

David Lewis was a philosopher and logician. His book *Convention. A Philosophical Study* ([Lewis(1969)]) proposes an extensive study of conventions as a scientific concept. The starting point of Lewis' study is the claim made by the philosopher W.V. Quine that language cannot be a convention since conventions need language to be established. Lewis' aim was to demonstrate that conventions may be established without any kind of communication device, and therefore that language could be conventional.

Lewis defines conventions as devices to solve coordination problems. A convention is a "coordination equilibrium" in a game where there are at least two coordination equilibria.[2] The simplest, rough, definition of convention Lewis gives is the following (p. 42):

> A regularity R in the behavior of members of a population P when they are agents in a recurring situation S is a *convention* if and only if, in any instance of S among members of P,
> 1) everyone conforms to R;
> 2) everyone expects everyone else to conform to R;
> 3) everyone prefers to conform to R on condition that the others do, since S is a coordination problem and uniform conformity to R is a (proper) coordination equilibrium in S.

Now, the following question arises: since a convention is an equilibrium in a (coordination) game where there are at least two equilibria, how can any member of P be sure of the regularity which will be followed by other members of P? This is the very same problem of coordination studied by Schelling and in fact Lewis produces a similar solution. Lewis sees that the resolution of any coordination problem involves the generation of *higher-order expectations* by individuals. That is, starting from the knowledge of individuals' rationality and preferences, each person in a coordination problem will draw expectations on what others expect on what everyone expects, and so on. More formally, define A as a bundle of information on the rationality, preferences and action of each member of P; a first-order expectation "I expect A" is produced by higher-order expectations of the kind "I expect that you expect A", "I expect that you expect that I expect A", "I expect that you expect that I expect that you expect A", and so on. If we note $R^n(A)$ as for "I expect A" with n the n-order of expectations, then the decision to conform to R as the result of the first-order expectation R(A) is produced by an infinite chain of higher-order expectations with n tending to infinity. Of course, as Lewis notes, humans hardly have the capacity to generate such an infinite chain of expectations. More often than less, expectations beyond a certain order are unnecessary, but the more

[2] Lewis defines a coordination equilibrium "as a combination in which no one would have been better off had *any one* agent alone acted otherwise, either himself or someone else" (p. 14). This definition is restrictive and applies only to coordination games. More recent studies of conventions drawing on Lewis' work adopt a more liberal definition. For example, Robert Sugden ([Sugden(2005)]) defines convention as any equilibrium in a game with at least two (Nash) equilibria.

expectations are of a high order, the more one has a reason to believe something and to act in a certain way (see [Lewis(1969)], p. 33). One of the main achievements of Lewis' study was to demonstrate how such a chain of expectations might be generated to make coordination possible.

Lewis is the first to have proposed a theory of common knowledge.[3] The aim is to determine necessary and sufficient conditions for a state of affairs A to become common knowledge in a population. Lewis informally states those conditions as follows ([Lewis(1969)], p. 52-53): provided that everybody share the same background information and the same inductive standards, a state of affair A can make a proposition P common knowledge if 1) each of us has reason to believe that A holds, 2) A indicates to everyone that each of us has reason to believe that A holds, 3) A indicates to everyone that P. The proposition "A indicates P" means that if I have reason to believe that A holds, then I can reasonably believe that P is true. A and P are not linked deductively; rather, the link relies on inductive standard and some background information. Lewis demonstrates that these premises, taken together with reasonable premises on the rationality, inductive standards and background information are sufficient to generate the infinite chain of higher-order expectations.[4] It is interesting to formalize the epistemic conditions necessary for the higher-order expectations to be generated. For all persons i, j in a given population, a proposition P is common knowledge at a state $\omega \in \Omega$, if and only if for a state of affair A such that $\omega \in$ A, the following four conditions are satisfied:

1) $\omega \in K_i(A)$
2) $\omega \in K_i[K_j(A)]$
3) $K_i(A) \Rightarrow K_i(P)$
4) $K_i(A) \wedge K_i[K_j(A)] \Rightarrow K_i[K_j(P)]$

These four conditions read as follow[5]: the first condition states that each person i knows

[3]Robert Aumann[Aumann(1987)] has famously developed a model of common knowledge which is now considered as the standard understanding of the concept of common knowledge. However, Lewis' own theory significantly differs from Aumann's on several grounds ([Cubitt and Sugden(2003)]). The main difference is that Aumann uses a Bayesian framework. Hence, common knowledge is strictly a matter of knowledge, while for Lewis it is a matter of "reason to believe".

[4]In fact, if everyone shares the same inductive standards and background information, then if A indicates P to me, A must indicate P to you. Therefore, given premises 2 and 3, A indicates to both of us that each of us knows P. Now, combining premise 2 with this last proposition, I know that you know that I know P. Again, combining premise 2 with this last proposition gives 'I know that you know that I know that you know P', and so forth.

[5]See [Cubitt and Sugden(2003)] and [Vanderschraaf(1998)] for more detailed accounts on the epistemic assumptions of Lewis' theory of common knowledge. In our notation, we use the knowledge operator K as it is understood in standard Bayesian decision theory. As a result we conflate "to know that" with "to have reason to believe that". [Cubitt and Sugden(2003)] judiciously adopt a more precise formulation in terms of "reason to believe", emphasizing that according to Lewis higher-order expectations generating common knowledge rely on inductive reasoning. Having this point in mind, we use a simpler notation than theirs.

(has reason to believe) that A holds at the actual state of the world; condition 2 indicates that if A holds, i knows that j knows that A holds. Condition 3 is the formal statement of Lewis' third premise. The fourth condition states that person i can infer from his knowledge of the publicness of A that j knows P. This proposition makes explicit a crucial assumption in Lewis' theory: individuals are symmetric reasoners, that is they use the same kind of inductive reasoning. The two first conditions are purely empirical: the fact that A happened may or may not lead me to know that it happened and to know that you know that it happened. Together, they imply that A is a *public* event. Conditions 3 and 4 are epistemic: clearly, they necessitate that everyone uses similar reasoning and that everyone is sufficiently confident in this fact. If these four conditions are satisfied, then the following general proposition is true for all persons j:[6]

$$\mathrm{P(A)} = A \subseteq K_j(P)$$

P(A) is a proposition stating that at ω, an event A makes a proposition P common knowledge in a given population.[7] When P(A) is true, A is said to be an *indicator* or a *reflexive common indicator* of P and can be considered to be the result of a *framing effect* ([Gintis(2009)], p. 141): the indicator A frames the proposition P in a given population. Following Schelling, Lewis uses the salience of an event A as an explanation of the fact that A rather than any other event A' will make P common knowledge. Therefore, salience plays in Lewis' theory of common knowledge a similar role than in modern game theory: given the multiplicity of events that can be noticed by individuals, the salience of an event is a way to go beyond indeterminacy. Similarly to an explicit agreement, the salience of an event or an outcome makes people knowing (or believing) that everyone will behave according to what this event indicates. Hence, the salience of an event is defined by the property of this event to make a proposition or a set of expectations common knowledge.[8] Lewis essentially puts emphasis on the force of precedent as a special kind of salience allowing one to expect P from A: given reasonable inductive standard, the fact that A has led to P in the past leads one to reasonably expect that P will result from A next time. However, as Lewis himself recognizes, the same reasoning is valuable for any kind of salience.

Tacit but essential in Lewis' account is the link between salience and symmetric reasoning. As condition 4 formally states, for a proposition to become common knowledge the members of a population must be symmetric reasoners. This means that a salient event or a salient convention must be an event for which it is obvious (no matter how

[6] See [Gintis(2009)], p. 142, theorem 7.2. Obviously, given conditions 1 to 4, we have $\omega \in K_i[K_j(P)]$. Thus, we have $\omega \in K_j(P)$. Since $\omega \in A$, this proves proposition P(A).

[7] For any persons i, j, k, and given conditions 1-4, we have $\omega \in K_i[K_j[K_k(P)]]$, which gives $A \subseteq K_k(P)$. The same is true for any level of mutual knowledge, thus proving common knowledge.

[8] Following Lewis, it is essential to distinguish between the event proper and what the event announces or proposes. Since an event obviously always indicates itself, a public event is necessarily common knowledge. However, what makes an event especially salient is that people can *infer* something from it that it is common knowledge among them.

this obviousness might be explained) that it will indicate the same proposition to everyone.[9] Lewis takes symmetric reasoning as a reasonable assumption but does not give the necessary or sufficient conditions for it to be verified. If we follow Lewis' path, then any explanation of salience should ask how and when symmetric reasoning can obtain. Lewis seems to essentially consider symmetric reasoning as the sharing of obvious or 'natural' forms of reasoning. To take an example given by Lewis himself, imagine I am living in the city of Oberlin, Ohio at a time when local telephone calls were cut off after three minutes. A possible convention might be 'if I am the caller, I call back; if I am the called party, I wait that the original caller calls back'. If this conventional regularity has been followed for many years, then it will appear salient to any Oberlin residents and everyone will expect that when a call is cut off the caller calls back. Indeed, this expectation will be common knowledge. The obviousness of this convention lies in the fact that it is a recurring practice and that it seems *natural* to expect that what has worked in the past should work in the future. Clearly, the precedent makes this solution salient because individuals are deemed to share a common inductive form of reasoning where past events are taken as a prediction of what future events will be. Hence, this 'intuitive' form of reasoning gives to the precedent a particular salience. It appears that salience because of precedent is a particular case of *natural salience*. In the rest of the paper, we will defend another conception of salience, 'public salience'. We interpret symmetric reasoning as a matter of interpretation or of common understanding. We argue that the condition of symmetric reasoning has institutional causes and is a product of a tacit collective agreement.

3 Salience and Correlated Equilibrium

Because Lewis tries to explain how a given fact or proposition x becomes commonly known in a population, it provides the epistemic foundations for a complete understanding of salience. But Lewis' theory remains incomplete because it does not explain *why* a specific event A became an indicator while a similar but not identical event A' fails to generate common knowledge of expectations in the same population. Indeed, salience is still an *explanan* and Lewis did not attempt to provide an *explanation* of it. Yet, Lewis' theory of common knowledge gives a critical insight by indicating that salience operates through the fact that individuals are symmetric reasoners. In this section, we use the concepts of correlated equilibrium and correlated strategies to argue that a correlated device will be used because of its salience. The salience of these devices is a function of preexisting institutions.

It is useful to think of salience as a particular kind of correlating device since it operates

[9] If individuals are not symmetric reasoners, then an event A cannot make a proposition P common knowledge. Therefore, A cannot be salient because nobody will have reason to expect that others will conform to P. It is obvious that the condition of symmetric reasoning is needed to make salience a way to overcome indeterminacy. To give a convincing account of symmetric reasoning is then crucial.

by correlating each players' conjectures regarding what others will do. Furthermore, this makes clearer our point that salience cannot be reduced to individual (or 'natural') properties alone. Hence, salience can be understood through the concept of *correlated equilibrium*. An interesting example is the emergence of a convention of property in the "hawk-dove" game ([Bowles(2006)], [Sugden(2005)]):

	Hawk	Dove
Hawk	(V-C)/2 ; (V-C)/2	V ; 0
Dove	0 ; V	V/2 ; V/2

Two players bargain for an asset V and can fight ("Hawk") or negotiate pacifically ("Dove"). If they both fight, they win the asset with a probability of one half and loss C the rest of time. If they both negotiate, they share equally the asset. If one fights while the other tries to negotiate, the hawkish player wins the prize. With C > V, we have a game with two Nash equilibria in pure strategies: [Hawk; Dove] and [Dove; Hawk] and an equilibrium in mixed strategies where each player plays "hawk" with probability equal to V/C. Under an evolutionary setting, and under any plausible evolutionary dynamic, the population will settle around this last equilibrium, with a proportion V/C of the population playing "Hawk" and a fraction of 1-V/C playing "Dove". It is easy to see that this equilibrium is inefficient since then hawk players will meet with a probability of $(V/C)^2$, leading to a global loss of C in each of these encounters. However, we can assume that at each interaction one of the two players physically possesses the asset; then, each player is able to define himself and his opponent respectively as "possessor" and "challenger". That is, contextual features of the interaction make possible the labeling of players. Such labeling changes the game since strategies can now be made contingent to the role (possessor or challenger) of each player. In fact, strategies such as "if possessor, always play hawk, if incumbent always play dove" and "if possessor, always plays dove, if challenger always plays hawk" are *correlated equilibria* ([Aumann(1987)], [Bowles(2006)], [Gintis(2009)], [Smith(1982)], [Sugden(2005)]).[10] It is easy to see that both correlated strategies are evolutionary stable (*i.e.* they are best response to themselves) while the uncorrelated mixed-strategy, while still an equilibrium, is no longer evolutionary stable ([Sugden(2005)]).

The correlated device (here, the physical possession of the asset) is not part of the mathematical description of the basic game. It is still possible to rewrite the game by adding the correlated strategies and defining the added outcomes. The evolutionary biologist John Maynard Smith ([Smith(1982)]) dubbed the correlated strategy "hawk if possessor, dove if challenger" the "bourgeois strategy" since it leads to the emergence of a norm of property. Interestingly, the converse strategy is also an equilibrium but, except for one

[10] A correlated equilibrium is an equilibrium of correlated strategies, that is strategies which are contingent of a random signal. A correlated equilibrium corresponds to a Nash equilibrium in the underlying game with the correlated strategy available in the set of strategies of each player.

species of spiders mentioned in ([Smith(1982)]), this equilibrium is hardly found in nature and in human societies. Undoubtedly, correlated strategies necessitate both labeling and pattern recognition: following Brian Skyrms' expression ([Skyrms(1996)]), correlation results from the breaking of the symmetry between players through labeling. Once players are able to distinguish between different roles in the interaction, they have then to be able to find patterns of behavior associated with different roles ([Sugden(1998)]). This begs a fundamental question: what are the reasons making sense of the pervasiveness of some correlated strategies such as the bourgeois strategy in nature and human societies while other correlated strategies are virtually non-existent in spite of their equilibrium property? In other words, where does the salience of some correlated strategies come from? Since at any moment in any interaction, a potentially infinite number of devices can help the players to correlate their expectations, the problem is to determine what are the properties making some of them able to generate common knowledge of expectations in a population. Though sometimes some natural properties can be an explanation, we contend that in many cases the salience of a correlating device is better explained by its social properties and the institutional environment of the players.

4 A Short Digression: Where Do Common Priors Come From?

It is attractive to interpret salience as a special form of correlated equilibrium because it helps to make a link between salience and the so-called 'common priors' assumption briefly discussed in section 1. This assumption holds that rational agents draw the same inferences on how a game is to be played. Meanwhile, a salient outcome or a salient strategy is one which everyone expects to be played. Hence, to assume that a strategy is played because it is salient is the same has to assume that players have common priors on the way the game will be played. The association between the common priors assumption and the concept of correlated equilibrium is due to an important theorem first discovered by Robert Aumann ([Aumann(1987)]) that establishes that Bayesian rational players with common priors in any epistemic game G will implement a correlated equilibrium in the corresponding game G* augmented by an initial move made by Nature (see also [Gintis(2009)], pp. 136-139).

An epistemic game G is a normal form game with a set $N = (1, ..., n)$ of n players, a set $S = (S_1 \times ... \times S_n)$ of pure strategies and with a function $u_i : S- > \mathbf{R}$ mapping any combination of strategies into a n-tuple of real numbers. In addition, G includes a set Ω of possible states of the world, a knowledge partition P_i of Ω for each player and finally a subjective prior $p_i(.;\omega)$ that defines each player's beliefs regarding how the game will be played for each state of the world $\omega \in \Omega$ ([Gintis(2009)], pp. 83-84). A player's subjective prior states his beliefs concerning the actual state of the game when the actual state of the world is ω. Each state ω specifies the strategy profile that will be used in the game. Hence, a player's subjective prior defines his beliefs over the strategy profile that will be played by others. For a given state of the world ω, a Bayesian rational

player will choose the strategy that maximizes his expected payoffs given his conjectures on how the other players will play. To say that players have common priors over Ω is simply to say that they have the same beliefs distribution over each possible state of the world ω. Aumann's theorem then demonstrates that when Bayesian rational players have common priors over Ω, they will implement a strategy profile that corresponds to a correlated equilibrium. The converse is also true: any correlated equilibrium in a game G* corresponds to an epistemic game G where players have common priors such that for each state of the world ω, it is rational for them to play their part in the correlated equilibrium ([Gintis(2009)], p. 136).

Regarding the working of salience, the significance of this theorem is great. We have suggested above that salience works as a correlating device: an exogenous feature helps the players to make convergent and consistent expectations regarding what each of them will do. Salience allows players to form their expectations in a non-independent way. Aumann's theorem demonstrates that it is the same as to assume that (Bayesian rational) players have common priors over Ω even when there is no explicit correlating device. Hence, the problem of the origins of salience can be at least partially answered by determining from where do common priors come.

The assumption of common priors solves the problem of indeterminacy that occurs when two or more equilibria exist in a game. In fact, to assume common priors is the same as to assume that players have already agreed on the way they will play.[11] If this agreement corresponds to an equilibrium, then nobody has interest to deviate provided that others implement the agreement. Interestingly, the common priors assumption seems to lead unavoidably from a departure to any strict individualistic understanding of rationality in a game because it seems difficult to see how multiple individuals can entertain the same subjective prior without any kind of intersubjective relationship. However, economists have originally tried to justify this assumption without making any appeal to some form of 'collective' rationality. For instance, John Harsanyi has argued that two rational individuals submitted to the same information must necessarily and independently come to the same conclusion. Robert Aumann ([Aumann(1976)]), who dubbed this claim the 'Harsanyi doctrine', noted that this is the same as to assume that rational individual have common priors. The Harsanyi doctrine implies that any disagreement between two rational individuals originates necessarily in the differences between the information on which they rely. Pushing Harsanyi's claim farther, Aumann ([Aumann(1976)] demonstrated that if rational individuals have common priors and if their posterior beliefs are common knowledge, then these players cannot 'agree to disagree': because of the Harsanyi doctrine, any disagreement must result from the fact that at least one individual has private information; since the disagreement makes common knowledge individuals' beliefs,

[11] The assumption of common priors is usual in sequential games with incomplete information. When there is uncertainty regarding the type of one or several players, it is common to assume that the actual types are determined by a random move made by Nature and that the players know the relevant probabilities for each type to be picked.

it reveals that at least one of them possesses some private information. If individuals are Bayesian rational, they will revise their beliefs accordingly until they reach an agreement.

This last theorem *assumes* that individuals have common priors. The fact that disagreements over many topics in modern societies are pervasive seems to indicate that either individuals do not have common priors or that their posterior beliefs fail to be common knowledge. Indeed, the Harsanyi doctrine seems to apply only in a narrow set of circumstances. We can expect individuals to have common priors only for straightforward and recurrent problems. In particular, people probably have common priors regarding several *natural events* such as for instance the probability that an earthquake occurs in a given region. People probably also entertain similar beliefs regarding the probability that a person jumping from the top of a building survives. These types of events involve basic physic laws or obvious consequences such it is reasonable to expect that all persons endowed with normal mental faculties will hold similar beliefs. Matters are less clear for very rare natural events and for many situations where individuals have to form beliefs over others' intentions or expectations. Here, to assume common priors seems heroic, unless we can establish that in one way or another individuals have been able to tacitly agree over consistent and convergent beliefs.

The similarity between the Harsanyi doctrine and the assumption of symmetric reasoning (section 2) is striking. Indeed, the Harsanyi doctrine states two rational individuals reason in an identical way. In other words, it is as to assume that everyone infers the same conclusion from the same set of premises. The Harsanyi doctrine and the assumption of symmetric reasoning pose no problem regarding a small set of natural or basic events. The 'natural' salience of some phenomena falls in that category: for example, people have the ability to uncover some natural patterns in a list of numbers (and everyone can expect everyone to have this ability) because these patterns follow both from a basic knowledge of arithmetic and from basic human mental capacities. As we have noted however in section 2, symmetric reasoning in many social interactions calls for an explanation. Similarly, the assumption of common priors on which the Harsanyi doctrine is grounded is all but straightforward in complex social interactions. The rest of the paper connects the assumptions of symmetric reasoning and of common priors with the fact that individuals share a common institutional heritage.

5 Institutions and Salience

Indeed, we can try to explain salience by focusing on the social mechanisms responsible for it, in particular institutions. Institutions can be defined as rules, norms, beliefs and organizations that together produce a regularity of behavior ([Greif(2006)]). In a game-theoretic framework, institutions will play a twofold role: first, institutions set the rules of the game, that is they define the structural features of an interaction. Once

we model an interaction as a game, the number and the identity of players, the set of strategies of each player and the consequence function defining gains for each possible outcome, are all defined by preexisting institutions. Second, we argue that institutions also contribute to define how players frame a situation; institutions are not solely objective features, there effect expands at the intersubjective level: they influence beliefs and expectations.[12]

To have an intuition of this idea, take the following example. In a recent paper reviewing the contribution of neuroeconomics and of behavioral economics, David Levine ([Levine(2011)]) argues that rational behavior can cause financial bubbles. Therefore, he estimates that we do not need to look at people's psychology and brain functioning since a theory of irrational behavior is unnecessary. He gives a sensible illustration (p. 17):

> The situation in a market panic is similar. Suppose you turn on the television and notice the Chairman of the Federal Reserve Board giving a speech indicating that the financial sector is close to meltdown. It occurs to you that when this happens, stocks will not have much value. Naturally you wish to sell your stocks – and to do so before they fall in price, which is to say, to sell before everyone else can rush to sell. So there is a "panic" as everyone rushes to sell. Individual behavior here is rational.

The market panic situation Levine is describing has the form of an assurance game (payoffs are those of the row player, in column are represented all the other players):

	Sell	Do not sell
Sell	5	6
Do not sell	0	10

In this situation, everyone will be better off by not selling his stocks but if others decide to sell, the best one can do is to sell. That everyone rationally chooses to sell after having seen on television the Chairman of the Federal Reserve predicting a financial meltdown is another instance of correlated equilibrium: the speech of the Chairman on television is a public signal allowing correlated strategies. It indicates to each player to sell his stocks and clearly, each player has to follow the signal provided that everyone else does the same: the correlated strategy "sell if the Chairman announces a financial meltdown" is a best reply to itself and so corresponds to a correlated equilibrium. Levine gives this example to make the case of his argument that looking at psychological mechanisms is unnecessary. However, actually both psychological and institutional factors enter the picture. The formalism of Lewis' theory of common knowledge is useful here. The Chairman's speech is a public event A: the fact that the Chairman gives

[12]The idea that institutions act on beliefs and have an intersubjective nature can be found in both economics and philosophy. For the former, see [Aoki(2001)], [Greif(2006)]; for the latter, see [Searle(1995)].

his speech at television leads everyone to reasonably believe that everyone knows A. The public announce of the Chairman leads each individual to believe the proposition P that a financial meltdown is coming. Provided that each individual are symmetric reasoners, everyone knows that everyone knows P: that a financial meltdown is coming becomes common knowledge. All four conditions for common knowledge to be generated are satisfied. Then, the Chairman's speech A is a reflexive common indicator making everyone's expectations converging to P, making P a self-fulfilling prophecy.

In this example, the Chairman's speech makes the financial meltdown a salient possibility because of its framing effect. What we want to know are the causal mechanisms lying behind this framing process. There are clearly biological and psychological mechanisms: the observation of the event A is possible because humans have some cognitive and sensory faculties. Moreover, the way each individual links A to P is based on reasoning faculties, both deductive and inductive. Finally, the assumption that we are symmetric reasoners is partly grounded on the evolutionary history of the Human species. Therefore, framing is partially tied to proximate (psychological) mechanisms. But it is clear in this example that we cannot understand how the event A makes the proposition P common knowledge if we ignore the institutional context that is tacitly assumed. In fact, this example is full of institutional signs that make the example understandable and meaningful for the community of economists and beyond. For instance, one cannot understand how the Chairman's speech becomes an indicator if one ignores the institutional role of the Federal Reserve and of its Chairman in the US economy. Indeed, it is probable that a similar announce made on television, on the same day at the same hour, by any other individual (even if she is an economist) would not make the proposition P common knowledge. Now, it is doubtful that the propensity of A to make P common knowledge is tied to the personality of the Chairman. It is more likely that this propensity is best explained by the *institutional role* that is occupied by that person. That the head of the Federal Reserve is an important person susceptible to have influence on the economy and that has access to important information is the product of the institutional structure of the American economy. That every American adult person knows this information can be explained by the fact that this is part of a common culture which is itself generated by the fact that Americans live in the middle of the same economic and non-economic institutions. In a nutshell, salience through framing results from a shared *institutional heritage*. We expand on this point in the next section, introducing the concepts of embeddedness and of common understanding.

6 Salience as a Public Phenomenon

We have just argued that institutional factors are important to understand how salience works. Institutions lead people to frame significantly a given situation in the same way, so that a thing someone views as salient is also seen as salient by others. Starting from

Schelling's account and as it is clear in the game-theoretic framework of Lewis, salience occurs in *strategic* interactions. Therefore, it is not sufficient to see something as salient to choose it, we have also to expect that everyone see the situation as we see it. Therefore, salience has an intersubjective dimension and so implies some form of *collective reasoning*. In this section, following [Postema(2008)], we define salience as a public phenomenon.

That salience has a public (or collective) dimension is already featured in Lewis' theory of common knowledge: the conspicuousness of any solution builds on the assumption that individuals share the same background information and similar inductive reasoning modes. We can go further by pointing that individuals will form convergent expectations not only if they share modes of reasoning and information but moreover *if they recognized this fact*. It makes salience eminently social because only members of a specie endowed with the ability to create culture can develop reflexively such a knowledge. As Gerald Postema puts it ([Postema(2008)], p. 45): "Salience, so understood, is not merely a matter of psychological fact, it is social fact - or rather, it is a fact available only to beings capable of exercising a certain kind of social capacity".

In an important series of articles on Lewis' approach of conventions, the philosopher Margaret Gilbert ([Gilbert(1989)], [Gilbert(1990)]) argues that the salience (especially because of precedent) of a particular action or outcome does not in itself give one sufficient reasons to expect that others will choose this action or outcome. Indeed, to play the salient strategy, it seems that the same *reductio ad infinitum* thinking is necessary than in any coordination game with multiple equilibria: I will choose the salient strategy S if I expect you to choose S; you will choose S if you expect me to expect you to choose S, and so on. Discussing the specific case of salience because of precedent, Gilbert speaks of the 'impotence of the precedent' to emphasize that precedent cannot be a device helping to choose.[13] This impotence makes salience useless for rational players to coordinate in strategic interactions because what has occurred in past interactions provides no rational reason to expect that it will continue to occur. Therefore, if salience helps individuals to coordinate, it must be on the ground of non-rational reasonings. The charge that salience is useless in coordination problems might be answered in two different ways. One can argue that salience evolves through inductive learning rules. Inductive rules do not have any rational basis in a pure deductive sense. They are rules of thumb which are pragmatically efficient. Hence, even if salience as a choosing device is non-rational (as Gilbert argues), it is still pragmatically justified for evolutionary reasons. The problem with this line of argument is that it is difficult to determinate what the actual inductive rules responsible for the evolution of salience truly are. Moreover, such evolutionary explanations are difficult to interpret.

[13] "[T]he fact that two agents have in the past done their part in an equilibrium so far says nothing about what either will do in the future" ([Gilbert(1990)], p. 12.)

A second answer consists in given up completely the concept of salient. This is what Gilbert suggests when she proposes her "Group Principle model" as a substitute. Gilbert's basic idea is that rational individuals are totally justified in adopting a principle of action covering all situations of a similar kind. The group principle model states that the members of a population are likely to jointly accept a specific principle; once such a collective principle has been defined, no rational individuals will have interest in adopting a different personal principle. The joint acceptance of a principle of action implies a form of collective agreement which may be either verbal or tacit. The agreement on the use of a particular principle of action between members of a population seems analogous to Lewis' requirement that individuals must be symmetric reasoners for a proposition to become common knowledge. Crucially, however, it indicates that symmetric reasoning is not a natural feature but quite the contrary the collective product of an agreement among members of a population. Hence, if as it has been suggested above (see sections 2 and 3), salience implies symmetric reasoning, that means that we can directly link salience phenomena with some kind of collective agreement.

We contend that the only way to make sense of the seemingly natural attraction exerts by specific salient events is that individuals share social capacities and characteristics. Postema ([Postema(2008)]) points out the importance of 'common practical reasoning'. Common practical reasoning leads people to make sense in a harmonic way of social features without complete deliberation.[14] But social capacities and characteristics are more than common practical reasoning. In fact, reasoning needs inputs insuring each person that everyone understands the situation in the same way. This implies that salience is public because it cannot resume to the private description of the situation of each individual. More than common reasoning, salience necessitates what we will call common reflexivity or more precisely, *high-order common reflexivity*. Below, we show that high-order common reflexivity allows individuals to be confident in the fact that they have a common understanding of a given situation and that the higher the degree of common reflexivity is, the more we approximate a form of collective thinking.

7 Common Understanding and Collective Thinking

Reflexivity is a circular form of reasoning where the thinker has conscience of and reflects on his own thinking process. The famous quotation of the French philosopher René Descartes "*je pense donc je suis*" points out the fact that reflexivity is at the heart of what defines any human being. Arguably, reflexivity is what distinguishes humans from any other intelligent animal. Conventional game theory does not ignore the importance of reflexivity. Indeed, the assumption of common knowledge of rationality (not in Lewis'

[14]Postema builds his argument on the example of jazz music. Jazz playing is largely grounded on improvisation and relies on intuitive, implicit thinking. Despite of this, jazz thinking appeals to capacities that when exert produce a harmonious, complex and articulated form of social knowledge.

sense) ensures that every individual has the same cognitive and reasoning abilities such as when a player asks "what will *he* do?", it is as if he would ask "what will I do?". This is precisely this kind of reflexivity that generates some of the puzzles presented in the first section.

In a typical Cartesian understanding, reflexivity is an individually-centered thinking process; it only involves the first person of the singular. Reflexivity helps the individual to define himself and to understand the implications of his own thinking activity. However, as argued above, salience is public and social. It can be understood only by abandoning the subjective point of view to adopt the *inter*subjectivity perspective. That is, we have to characterize the fact that individuals share social capacities (modes of reasoning and background information) and that individuals can figure this. Indeed, inside a population, individuals will be able to use salience as a choosing device if they are confident that their conception of salience is shared by everyone. For any population P and any set K of background information and modes of reasoning, we define the concept of *common reflexivity* as follows:

> *Common reflexivity*: there is common reflexivity $CR_P[K : S- > C]$ in a population P if the members of P use the same set K of background information and practical reasoning such that, from a set of premises S, they reach the same conclusion C.

Therefore, common reflexivity implies the symmetric reasoning condition found in Lewis' theory of common knowledge. It is also similar to the statement made by the Harsanyi doctrine: individuals using the same information must come to the same conclusion, provided that they have common priors. Common reflexivity indicates that every member of the population P will interpret a given situation (defined by S) similarly. Moreover, the fact that humans are endowed with properties of reflexivity means that they can acknowledge that they share some forms of reasoning with others. This will generate a *higher-order common reflexivity*:

> *High-order common reflexivity*: we say that there is second-order common reflexivity when every member i of a population P uses the fact that there is common reflexivity in P has an input in the set of background information K. Formally, $CR_P^2 = [CR_p \in K]$. Generalizing, there is nth order of common reflexivity when each member of P uses the fact that there is (n-1)th order of common reflexivity in P has an input in the set K. Formally, $CR_P^n = [CR_p^{n-1} \in K]$.

A high-order of common reflexivity in a community implies that everyone acknowledges that everyone interprets the information contains in K identically. As in Lewis' theory of common knowledge, infinite-order common reflexivity is both non-realistic and unnecessary: it is non-realistic for clearly evident cognitive limitations (even if this kind of

reflexivity is largely unconscious) and it is unnecessary because the marginal confidence brings by a supplementary level of order of common reflexivity is decreasing as the level of order is higher. We are able now to define the concept of *common understanding* as the theoretical limit of infinitely high-order of common reflexivity:

> *Common understanding (CU)*: there is *common understanding* in a population P when the degree n of common reflexivity in P reaches an arbitrarily high level.

CU means that every member of a community is confident in the fact that everyone interprets recurring or representative situations in the same way. Therefore, salience as a public phenomenon is the product of *CU*: starting from a situation defined by premises S, I will infer from K that I must do 'x', making 'x' salient because I know others will reach the same conclusion. Because K incorporates an embedded chain of n-1 orders of common reflexivity, I know that 'x' is similarly salient for others and I know that they know this, and so on. Common understanding explains why some correlated devices and not others are salient and generate correlated equilibria. In Levine's example of the Chairman's speech, the interpretation of the prediction of a financial meltdown on television is obvious to anyone is acquainted with market economies having a long history of financial turbulences. Arguably, every leaving people in market economies have a common understanding of the way these economies work and so "know" that a signal sent by the Chairman of the Federal Reserve is significant and meaningful to everyone.

It is interesting to note the analogy between the common understanding concept and Gilbert's joint acceptance of a principle requirement. Common understanding may be thought as the result of an agreement reached in a tacit collective bargaining. More generally, *CU* is an integral component of collective thinking and 'we-reasoning'.[15] [Hakli, Miller, and Tuomela(2010)] argue that any occurrence of we-reasoning is made of a three stages process resulting in situations amenable to a game theoretic description where every characteristic of the game is common knowledge. For any already constituted group with a set of collectively accepted goals[16], a we-reasoning process follows three stages in the following logical order:

WR1: the parties share and acquire information on each other, notably their preferences and their modes of reasoning. This stage ends when this information is made common knowledge among the parties.

[15] See [Gold and Sugden(2007)] for a survey of the main theories of collective thinking in economics and philosophy.

[16] Hakli et *al*. start from the premises that the group already exists and that that its members have at least already agreed on the goal to find out what each group member wants. These premises are reasonable given our conceptual problem, which is how individuals belonging to the same population are able to coordinate. Moreover, it is clear that this initial state must not necessarily be the product of a verbal/explicit agreement.

WR2: the parties form a *joint intention to act*, that is they agreed to choose an outcome in the commonly known matrix generated in the preceding stage.
WR3: the parties carry out this joint intention by actually choosing the outcome they have agreed on during the preceding stage.

Common understanding as we have defined it forms in the first two stages. In stage WR1, parties build the set K of background information and practical reasoning that they will share. In stage WR2, the set K allows members of the group to form a joint agreement on a specific outcome; hence, this agreement acquires a particular salience and gives valuable reasons for each individual to use it as a correlated device. Contrary to Gilbert, we retain salience as a useful feature helping individuals to choose in a strategic interaction. But, similarly to her 'Group Principle model', we advance the importance of a preexisting joint agreement. Of course, the process made of stages WR1 and WR2 and resulting in common understanding and a shared conception of salience will generally be purely fictitious. Finally, it should be clear that common understanding *does not imply* complete team reasoning and the disappearance of individual agency. Individuals may still behave according to their personal preferences through a reasoning of the kind "*we* interpret the situation as X, so I should do x".

It is useful to contrast this conception of 'public salience' with the 'natural salience' that is apparent in Lewis' theory of convention and also in game-theoretic formalizations of salience. According to our account, salience is not a natural fact that results from 'obvious' inductive rules that any individual may reasonably use. Salience is rather the product of a social 'agreement' revealing to every individual that he shares with other a similar understanding of the way the world works and about what matters. Using the same example of the phone call convention in the city of Oberlin, what makes the regularity "if I am the caller, I call back; if I am the called party, I wait" salient is first and foremost that I am a resident in Oberlin. This is not to say that precedent has no weight has a device to make coordination successful or that this particular solution has not a natural obviousness.[17] However, it might be possible that in another town, the residents have tacitly agreed on a different conception of salience. As an example, the solution "the person who was speaking when the call cuts off calls back" seems equally salient. The choice between these two salient solutions clearly requires a common understanding of the situation generated by the institutional environment.[18]

[17]Natural salience and public salience are not necessarily antagonistic. We do not dispute the fact that salience might sometimes have a purely natural (and so, individual) origin.

[18]Hence, our conception is similar to Verbeek's salience by convention indicating that salience is defined on conventional grounds ([Verbeek(2002)]).

8 Common Understanding as the Product of Embeddedness

Where does CU come from? How are higher-orders of common reflexivity generated? These are questions one has to answer to give a complete explanation of salience and focal points. As suggested in the fifth section with the example of the market panic following a speech of the Chairman of the Federal Reserve, institutions largely explained why this speech acquires salient properties and figures as a reflexive indicator. Because individuals in the same community share the same institutional heritage, they have confidence in the fact that they interpret the same institutional fact in a similar fashion. They have a common understanding of many situations and they are used to rely on the same institutional signals to make sense of what is happening. Therefore, this is this shared institutional heritage that brings common understanding into a population.

The sociological concept of *embeddedness* is helpful to explain the role played by institutions in generating common understanding and salience.[19] The sociologist Mark Granovetter ([Granovetter(1985)]) has proposed this concept as a remedy to the "over- and undersocialized conceptions of human action in sociology and economics". According to Granovetter, the oversocialized conception of human action is pervading in modern sociology. It developed as a response to the undersocialized conception of the individual one can found in economics but in fact originates in the political philosophy of Thomas Hobbes. Since Granovetter wrote his article in 1985, game theory has largely displaced the atomistic conception of the consumer and of the producer in economics. But Granovetter's article already points out the lack of contextual and historical dimension of more recent microeconomics and the rather mechanistic treatment of the interactions between individuals and collective structures. Granovetter's embeddedness concept emphasizes the importance of networks and personal relationships in and out of markets for generating trust and discouraging malfeasance ([Granovetter(1985)], p. 490). Because economic relations are mostly organized through network structures, agents actually are engaged in repeated interactions where the identity of partners is known. Reputation building is therefore of primary importance and leads to interactions where it is common knowledge that agents have common interests. Granovetter then goes on to argue that social structures, more than morality or institutional arrangements, are the main explanatory device to understand how trust, coordination and cooperation are generated.

Granovetter's thesis has been heavily discussed and has received some strong criticism, mainly coming from sociologists. To make embeddedness a useful concept to understand how common understanding and salience emerge in a population, at least two major amendments should be made to Granovetter's analysis. First, Granovetter opposes "social structures" and "institutional arrangements". Given our definition of the

[19] Initially, the concept of embeddedness has first been proposed by the Hungarian economist and anthropologist Karl Polanyi. However, Polanyi largely used this concept in a structuralist sense which largely departs from its use in modern sociology.

concept of institution, it should be clear that institutions are part of social structures. If we identify networks with social structures, we can go further: institutions define networks by imposing rules and norms regulating social interactions. Second, Granovetter's argument about trust building through personal relationships seems perfectly appealing from a game-theoretic point of view. In fact, repeated games have helped economists to build an important literature on reputation and learning in strategic interactions. While economic sociology can help to understand the building of network structures generating repeated interactions, embeddedness proper does not bring much insight on the generation of trust. However, keeping with the idea that embeddedness points to the fact that individual action occurs through network relations and more largely in communities defined by specific institutions, we have one way to explain how common understanding can be generated.

Reasonably, there are two general ways through which an arbitrarily high-order of common reflexivity can be generated among a population. These two ways are somewhat isomorphic to the possibilities for the joint agreement to be verbal (explicit) or tacit. A verbal joint agreement defining common understanding and salience implies an *actual* bargaining or conversation. Hence, a first possibility for common understanding to exist is through *repeated personal relationships*: repeated interactions between two or more individuals (implying eventually communication and cheap talk) make easier and faster the sharing of relevant information on the way each participant understand the situation. Clearly, embeddedness might be responsible for repeated interactions (particularly if we use Granovetter's definition).

A second, more general way for common understanding to occur is through a tacit joint agreement. Even if two individuals never directly interact with each other, they can still know that they share a common conception of salience because there are obvious signs that they belong to the same community or that they have a similar collective identity. More generally, if individuals share the same institutional heritage, then common understanding can obtain. As an example, even if I cannot be sure that everyone has seen the Chairman of the Federal Reserve on TV (the event A), I still can be confident in the fact that the proposition P that the event indicates is common knowledge because we share a common conception of salience giving a particular weight to every word the Chairman pronounces in every instance.Similarly, if you and I are French, then we can have confidence that any intervention of Nicolas Sarkozy on television will have a particular salience for both of us, even if we do not know each other personally. In both case, we know that a *class* of similar events will be salient to everyone because we have a common understanding of a number of situations. In the same vein, Michael Chwe ([Chwe(2001)]) provides numerous examples demonstrating how public rituals help to foster common knowledge among a population. He explicitly makes the link between the creation of common knowledge and the existence of communities. Though he does not use the concept, Chwe also emphasizes the importance of common understanding

of a situation or of a communication to make common knowledge possible: "The point here is that common knowledge depends crucially on how each person understands or interprets how other people understand or interpret a communication" (p. 83, see also p. 7). This requires that members of a community share a symbol system and a world view, making people symmetric reasoners. Then, to be recognized as a member of community can be thought to be recognized as a party in the tacit joint agreement.

Interestingly, our argument provides an alternative explanation to the 'analogy thesis' that Sugden ([Sugden(2005)]), following David Hume, endorses to explain the pervasiveness of some conceptions of salience, notably in the case of property conventions. Sugden's basic argument is that humans have a propensity to expand through analogy the frames used in a given situation to situations which are similar but different. This analogical reasoning makes salient some of the characteristics of an interaction I' because the participants take it to be analogous to an interaction I and use the same criteria of salience. Sugden's argument relies on a conception of salience as a natural phenomenon. However, according to our conception of public or institutional salience, individuals can use the same conception of salience not because they reason through analogy, but simply because as a members of a community they reason on the ground of the same tacit joined agreement; they use the same set of symbol system and the same world view.

What we have just argued is that salience is not a purely individual phenomenon. It is because salience is also public and social that institutions are part of it. Institutions make possible the common understanding necessary to overtake the "impotence of salience". However, institutions also contribute to salience in a more straightforward way. Indeed, institutions are salient precisely because they are institutions. Once an institution exists, it acquires a form of conspicuousness making it salient. However, any rule or convention must still be interpreted in a similar way by the members of a group to be effective.[20] Hence, the way an institution is interpreted depends on a shared conception of salience and of common understanding. Since institutions (note the plural) help to establish common understanding, this means that the way an institution is interpreted cannot be explained if we ignore the surrounding institutions that complement it.[21]

[20]The French philosopher Emile Durkheim was one of the first to state this point clearly when arguing that any contract is always incomplete. Even a rule or a contract on which everyone has explicitly agreed has to be interpreted because a rule cannot specify in all the required details the prescribed behavior. The same is true in the case of a tacit norm, but then the distinction between the norm 'as it is stated' and the norm 'as it is interpreted' is more conceptual than empirical.

[21]As an illustration, one can think of the way similar traffic rules (that is, institutions) are interpreted in different countries. While in some countries, everyone stops at red lights in other countries drivers use to interpret more loosely the same signal. In other words, red lights are always salient, but in different ways. For coordination to obtain, individuals must have a common understanding of the situation as a function of the country they are in.

9 Conclusion

This paper developed a conceptual and theoretical analysis establishing the institutional dimension of salience. Our aim was to show that the empirical significance of salience in economic interactions cannot be captured by a uniquely formal game-theoretic analysis but calls for a reflection on the institutional mechanisms underlying this phenomenon. Traditionally, game theorists have used two strategies to integrate salience in their formal analysis: either salience is formalized as the result (the explanandum) of an evolutionary process or salience is used as a quasi-primitive concept (an explanan) explaining the formation of norms and convention. The former strategy allows making salience endogenous by showing how individuals form convergent beliefs on the game they are playing. However, this strategy faces several difficulties, notably related to the empirical significance of the learning dynamics used in evolutionary models. The latter strategy shows how individuals achieve coordination but salience is an exogenous feature of the explanation. Moreover, because salience is not explained, it appears to be a natural phenomenon preexisting to individuals' interactions. In this paper, we have seek to provide an explanation of salience as a public phenomenon, building on and expanding David Lewis' theory of common knowledge.

Our study produced three significant results. First, we have established that salience is resolutely a public and collective phenomenon. Once we interpret salience as a form of correlated equilibrium, we have to include the public dimension to explain why some features and not others in an interaction act as a correlated device. Second, following Lewis, we pointed out that salience implies that individuals have to share the same background information and the same modes of reasoning. We developed the concepts of common reflexivity and common understanding, as parts of team reasoning, to capture this aspect. Finally, using the idea of embeddedness, we showed that institutions help to foster common understanding in a population.

This line of argument has implications for two important related topics. First, it asks how salience can be accounted for in formal analyses. Our analysis suggests that formal game-theoretic analysis should take salience as a primitive concept. But when game-theoretic models are used to develop historical and contextual analysis of the kind of [Aoki(2001)] or [Greif(2006)], our framework points to the importance of the historical narrative to ground the formal analysis in the historical, epistemic and institutional reality. Second, our analysis reinforces a statement made by others (*e.g.* [Gintis(2009)], [Hodgson(2007)]) on the limits of methodological individualism. If it is established that salience is needed to overcome the problem of indeterminacy in game theory, and if we accept our argument that salience cannot be reduced to an individual feature but has a public dimension, then we are lead to the conclusion that methodological individualism is a logical dead end.

References

[Aoki(2001)] Aoki, Masahiko. 2001. *Toward a Comparative Institutional Analysis.* The MIT Press.

[Aumann(1976)] Aumann, Robert. 1976. "Agreeing to Disagree." *Annals of Statistics* 4 (6):1236–1239.

[Aumann(1987)] Aumann, Robert J. 1987. "Correlated Equilibrium as an Expression of Bayesian Rationality." *Econometrica* 55 (1):pp. 1–18.

[Bacharach(1993)] Bacharach, Michael. 1993. "Variable Universe Games." In *Frontiers of Game Theory*, edited by Ken Binmore, Alan Kirman, and Piero Tani. The MIT Press.

[Bacharach(2006)] ———. 2006. *Beyond Individual Choice: Teams and Frames in Game Theory.* Princeton University Press.

[Bowles(2006)] Bowles, Samuel. 2006. *Microeconomics: Behavior, Institutions, and Evolution.* Princeton University Press.

[Chwe(2001)] Chwe, Michael Suk-Young. 2001. *Rational Ritual: Culture, Coordination, and Common Knowledge.* Princeton University Press.

[Cubitt and Sugden(2003)] Cubitt, Robin P. and Robert Sugden. 2003. "Common Knowledge, Salience and Convention: A Reconstruction of David Lewis' Game Theory." *Economics and Philosophy* 19 (2):175–210.

[Gilbert(1989)] Gilbert, Margaret. 1989. "Rationality and Salience." *Philosophical Studies* 57 (1):61–77.

[Gilbert(1990)] ———. 1990. "Rationality, Coordination, and Convention." *Synthese* 84 (1).

[Gintis(2009)] Gintis, Herbert. 2009. *The Bounds of Reason.* Princeton University Press.

[Gold and Sugden(2007)] Gold, Natalie and Robert Sugden. 2007. "Collective Intentions and Team Agency." *Journal of Philosophy* 104 (3):109–137.

[Granovetter(1985)] Granovetter, Mark. 1985. "Economic Action and Social Structure: The Problem of Embeddedness." *The American Journal of Sociology* 91 (3):pp. 481–510.

[Greif(2006)] Greif, Avner. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade.* Cambridge University Press.

[Hakli, Miller, and Tuomela(2010)] Hakli, Raul, Kaarlo Miller, and Raimo Tuomela. 2010. "Two Kinds of We-Reasoning." *Economics and Philosophy* 26 (03):291–320.

[Hargreaves-Heap and Varoufakis(2004)] Hargreaves-Heap, Shaun and Yanis Varoufakis. 2004. *Game Theory: A Critical Introduction.* Routledge.

[Harsanyi and Selten(1988)] Harsanyi, John and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games.* The MIT Press.

[Hodgson(2007)] Hodgson, Geoffrey M. 2007. "Meanings of Methodological Individualism." *Journal of Economic Methodology* 14 (2):211–226.

[Levine(2011)] Levine, David. 2011. "Neuroeconomics?" *working paper* .

[Lewis(1969)] Lewis, David. 1969. *Convention. A Philosophical Study.* Wiley-Blackwell.

[Luce and Raiffa(1957)] Luce, Duncan and Howard Raiffa. 1957. *Games and Decisions: Introduction and Critical Survey.* Dover Publications.

[Mehta, Starmer, and Sugden(1994a)] Mehta, Judith, Chris Starmer, and Robert Sugden. 1994a. "Focal points in pure coordination games: An experimental investigation." *Theory and Decision* 36:163–185.

[Mehta, Starmer, and Sugden(1994b)] ———. 1994b. "The Nature of Salience: An Experimental Investigation of Pure Coordination Games." *The American Economic Review* 84 (3):pp. 658–673.

[Postema(2008)] Postema, Gerald J. 2008. "Salience Reasoning." *Topoi* 27:41–55.

[Samuelson(2002)] Samuelson, Larry. 2002. "Evolution and Game Theory." *The Journal of Economic Perspectives* 16 (2):pp. 47–66.

[Schelling(1960)] Schelling, Thomas. 1960. *The Strategy of Conflict.* Harvard University Press.

[Searle(1995)] Searle, John. 1995. *The Construction of Social Reality.* Free Press.

[Skyrms(1996)] Skyrms, Brian. 1996. *Evolution of the Social Contract.* Cambridge University Press.

[Smith(1982)] Smith, John Maynard. 1982. *Evolution and the Theory of Games.* Cambridge University Press.

[Sugden(1991)] Sugden, Robert. 1991. "Rational Choice: A Survey of Contributions from Economics and Philosophy." *The Economic Journal* 101 (407):pp. 751–785.

[Sugden(1995)] ———. 1995. "A Theory of Focal Points." *The Economic Journal* 105 (430):pp. 533–550.

[Sugden(1998)] ———. 1998. "The role of inductive reasoning in the evolution of conventions." .

[Sugden(2001)] ———. 2001. "The evolutionary turn in game theory." *Journal of Economic Methodology* 8 (1):113–130.

[Sugden(2005)] ———. 2005. *The Economics of Rights, Cooperation and Welfare.* Palgrave Macmillan.

[Vanderschraaf(1998)] Vanderschraaf, Peter. 1998. "Knowledge, Equilibrium and Convention." *Erkenntnis* 49 (3):337–369.

[Verbeek(2002)] Verbeek, Bruno. 2002. *Moral Philosophy and Instrumental Rationality: an Essay on the Virtues of Cooperation.* Kluwer Academic Publishers.

[Young(1998)] Young, Peyton. 1998. *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions.* Princeton University Press.