Mindreading and Endogenous Beliefs in Games

Lauren Larrouy, Guilhem Lecouteux

MINDREADING AND ENDOGENOUS BELIEFS IN GAMES

Documents de travail GREDEG
GREDEG Working Papers Series

Lauren Larrouy
Guilhem Lecouteux

GREDEG WP No. 2017-01
https://ideas.repec.org/s/gre/wpaper.html

MINDREADING AND ENDOGENOUS BELIEFS IN GAMES

Lauren Larrouy, Guilhem Lecouteux

GREDEG Working Paper No. 2017-01

Lauren Larrouy: *Université Côte d'Azur, CNRS, GREDEG, France.* Postal address: CNRS-GREDEG, Campus Azur, 250, rue Albert Einstein, CS 10269, 06905 Sophia Antipolis Cedex, France. Email: lauren.larrouy@gredeg.cnrs.fr

Guilhem Lecouteux (corresponding author): *Université Côte d'Azur, CNRS, GREDEG, France.* Postal address: CNRS-GREDEG, Campus Azur, 250, rue Albert Einstein, CS 10269, 06905 Sophia Antipolis Cedex, France. Email: guilhem.lecouteux@unice.fr

Acknowledgements: we are grateful to Richard Arena, Cyril Hédoin, and two anonymous referees for helpful comments and suggestions. We also thank the participants to SING11 and to the 3rd International Conference Economic Philosophy, as well as seminar audiences at University of Bristol, Sciences Po Toulouse, and Université Côte d'Azur.

Mindreading and endogenous beliefs in games

Abstract (**100** words): we argue that a Bayesian explanation of strategic choices in games requires introducing a psychological theory of belief formation. We highlight that beliefs in epistemic game theory are derived from the actual choice of the players, and cannot therefore explain why Bayesian rational players should play the strategy they actually chose. We introduce the players' capacity of mindreading in a game theoretical framework with the simulation theory, and characterise the beliefs that Bayes rational players could endogenously form in games. We show in particular that those beliefs need not be ratifiable, and therefore that rational players can form action-dependent beliefs.

Keywords: prior beliefs; mindreading; simulation; action-dependent beliefs; choice under uncertainty.

Subject classification codes: B41, C72, D81.

Game theory can be defined as a mathematical theory of strategic interactions between rational decision-makers, involving sets of 'players', 'strategies' and 'payoffs'. The object of *classical* game theory is 'to propose *solutions* for games' (Sugden, 2001, p. 115, emphasis in original), and the two usual criteria for distinguishing between acceptable and non-acceptable solution concepts are (i) that the strategy profiles under consideration (the *equilibria* of the game) should be coherent with the rationality of the players, and (ii) that the set of equilibria should be relatively small, yet non empty (Sugden, 2001, p. 115). Bayesian decision theory – in the sense of Savage (1954), as the maximisation of subjective expected utility – was first introduced in game theory to discuss issues of incomplete information (Harsanyi, 1967-1968), while keeping the classical approach of studying solution concepts. It is only in the 1980s that game theorists started to analyse the Bayesian rationality of equilibrium play, by assuming that players maximise their expected utility, given some *subjective* beliefs about the actions of the others (unlike von Neumann and Morgenstern (1944, p. 19), who explicitly defined probability 'as frequency in the long run', and therefore as *objective* probability). Kadane and Larkey (1982a) were among the earliest[i] to suggest that choices in strategic interactions should be analysed through the lens of Bayesian decision theory. They argue that economists must investigate the 'question of where [prior beliefs] "come from"' (Kadane and Larkey, 1982a, p. 117), and suggest '[looking] to other disciplines such as cognitive psychology for predictive theories of decisional behavior' (Kadane and Larkey, 1982a, p. 118), rather than restricting the analysis to solution concepts.

Their position was however criticised by Harsanyi[ii] (1982a), who argues that the fundamental point of game theory is the study of rational choice *when the players expect each other to act rationally*: the object of game theory should therefore be to investigate 'how to choose these subjective probabilities in a *rational* manner' (Harsanyi 1982a, p. 123, our emphasis). Aumann (1987, p. 17) and Aumann and Drèze (2008, p. 81) raise a similar critique, and suggest that – rather than investigating the psychological foundations of prior beliefs – the mutual expectation of Bayesian rationality could be captured within the prior beliefs of the players through the assumption of common belief in rationality, combined with an assumption of common priors. Those assumptions – which are central in *epistemic* game theory (EGT) – seem however to be too cognitively demanding to provide a realistic account of choice in strategic interactions, but also raise some conceptual puzzles (see e.g. Morris, 1995; Gul, 1998; Levi, 1998; Bonnano and Nehring, 1999).

In line with Kadane and Larkey, the object of this paper is to offer a *psychological* theory of the formation of prior beliefs. However, we will also explicitly model how the players form their beliefs 'in a rational manner', i.e. when they believe that the other player is also a rational individual. We suggest in this paper introducing the players' capacity of *mindreading* – which is a widely investigated topic in cognitive sciences and in particular cognitive psychology and social psychology –, i.e. their psychological ability to attribute mental states to others and to form expectations about how they reason. Although different theories of mindreading have been suggested in the literature, the *simulation theory* (ST) of Goldman (2006) appears to be the most adapted for game theory, since it does not require more information than what is usually assumed in game theory (i.e. mutual knowledge of the game structure mainly). In addition to providing a theoretical explanation of the formation of players' prior beliefs in games, we highlight that endogenising the beliefs of the players also question the usual conflation between Bayes rational and *ratifiable* choices. An important implication is that the choice of Bayes rational players does not necessarily form a correlated equilibrium, and therefore that we need a new solution concept for a Bayesian analysis of games.

We begin by arguing that founding game theory on Bayesian rationality requires introducing a psychological theory of belief formation (section 1). We then introduce the simulation theory and describe how Bayes rational players form their beliefs in a rational manner through a 'massaging' process (section 2). We develop a game-theoretic formalisation of simulation and of the massaging process, and define a *subjective belief equilibrium* as the strategy profile resulting from the choice of Bayesian rational players, when they formed their beliefs by simulating the reasoning of the other players (section 3). We then show that simulation theory could justify the formation of *action-dependent beliefs*, and that Bayesian players can *always* rationalise the choice of a strategy profile Pareto-superior to a Nash equilibrium (section 4). Section 5 concludes by discussing some methodological implications of the multiplicity of subjective belief equilibria for a Bayesian analysis of games.

1. Where do beliefs come from?

We argue in this section that a proper Bayesian explanation of rational choices in games requires introducing a psychological theory of belief formation. We suggest that an operative theory of rational choice in games – i.e. a theory explaining how a player should rationally choose, given the nature of the game and of the other players – should:

(i) consider the problem faced by each player,

(ii) define a mechanism consistent with the rationality of the player that gives her a reason to play a specific strategy,

(iii) assess the final outcome, from our perspective as theorists.

The first step consists in describing the players' perception of the game, i.e. their beliefs about the types of the other players and about their strategies (for the sake of clarity, we will not consider here incomplete information games, and restrict our attention to the players' beliefs about the strategies of the others). The second step explains how the players choose, given their perception of the game (expected payoff maximisation, as implied by Bayes rationality, is a possible mechanism). The third and last step consists in defining a solution concept to characterise the resulting strategy profile.

As an illustration, consider the following Stag Hunt:

SH	A_2	B_2
A_1	(3;3)	(0;2)

B_1	(2;0)	(1;1)

From the perspective of classical game theory, several strategy profiles could be equilibria of the game: both A_1A_2 and B_1B_2 are Nash equilibria, A_1A_2 is the only strong Nash equilibrium, and B_1B_2 is the only stochastically stable equilibrium (Foster and Young, 1990). However, from your perspective as a player, classical game theory is of little help: if several strategy profiles can be equilibria of the game (depending on the solution concept), the theory does not give you any criterion to select one of these equilibria. Classical game theory cannot therefore offer an operative theory of rational choice in games, since it only considers whole strategy profiles, without providing a rationale for selecting an equilibrium (and hence cannot suggest a specific strategy for each player).

In line with our description of a theory of rational choice in games, 'EGT makes epistemic states of players an *input* of a game and devises solution concept that takes epistemics into account' (Brandenburger, 2010, p. 6, our emphasis). Your optimal strategy – and then the resulting equilibrium strategy profile – therefore depends on your beliefs about the choice of P2. But if P2 is also a rational individual (and thus faces the same indeterminacy problem than yours), can you treat her decision as a chance event? The difficulty we face here is how to define the prior beliefs of the players, knowing that those beliefs must be compatible with our mutual rationality (and therefore that they are *already* the result of some reasoning). We now highlight a central issue of EGT, namely that the prior beliefs supposed to capture the rationality of the players through the assumption of common belief in rationality are defined *ex post*, as a mere *representation* of the players' choice – they cannot therefore be treated as an input of the decision process.

While the terminology itself of 'prior beliefs' suggest that there exists an *actual* hypothetical prior stage from which, given the subsequent realisation of the state of nature, the players can update their beliefs – and then maximise their expected payoff given those 'posterior' beliefs – it appears that the priors have no genuine substantive meaning in EGT. Gul (1998) labels the former interpretation of priors as the 'prior view' and contrasts it with the 'hierarchy interpretation', in which case 'the prior are artifacts of a notational device to represent the infinite hierarchies of beliefs on [the players'] "posteriors" at the true state of nature' (Gul, 1998, p. 925). Aumann and Brandenburger (1995) explicitly endorse the latter interpretation, by emphasising that 'an interactive belief system […] does not suggest actions to the players. Rather, it is a formal framework – a language – for *talking* about actions, payoffs, and beliefs' (Aumann and Brandenburger, 1995, p. 1174, emphasis in original).

A major difficulty arises from this *behaviouristic* interpretation of payoffs. Heidl (2016, pp. 26-44) indeed suggests that preferences and payoffs can be interpreted either in a mentalistic or in a behaviouristic way. According to the mentalistic interpretation, 'preferences are understood as scientific refinements of the folk psychological concepts of desire and preference' (2016, p. 26), while the behaviouristic interpretation is that 'preferences are not mental entities but consistent patterns of choices' (2016, p. 27). The behaviouristic interpretation defines payoffs as von Neumann Morgenstern (thereafter vNM) utilities rather than material payoff[iii], meaning that the primitive of the game is the players' choices: the utility functions are defined *ex post*, as a *representation* of their choices. von Neumann and Morgenstern (1947) indeed showed that, if the choices of a player respect certain axioms, then it is *as if* the player was maximising an expected utility function. Similarly, Savage (1954), Anscombe and Aumann (1963), and Aumann and Drèze (2008) show that, if the player's choices respect certain formal conditions of consistency, then we can define a utility function and subjective beliefs such that the action that maximises the expected subjective utility of the player is precisely the choice we observed.

The behaviouristic interpretation raises serious methodological difficulties regarding the status of prior beliefs.[iv] Hausman (2012, pp. 28-33) indeed highlights that, if choice is jointly caused by preferences and beliefs, then we cannot simultaneously deduce preferences and beliefs from the choice of the players. If it were the case,

> the payoffs [in the behaviouristic interpretation] would say how individuals would choose. They would already incorporate the influence of belief, and belief could play no further role. If the revealed-preference theorist were right and payoffs already represented what strategy was chosen, there would be nothing left for game theory to do. (Hausman, 2000, pp. 111-112)

This means that the players' priors would be settled before the game, and the result of the other players' deliberation (i.e. their choice) would be comprised in the descriptions of the world provided by the prior stage.

For a given equilibrium strategy profile (e.g. a correlated equilibrium for Aumann 1987), we can build a prior belief such that the expected utility maximising choice of the players based on their posterior beliefs (which result from the realisation of the state of the world) correspond to this equilibrium strategy profile. Returning to the steps (i), (ii), and (iii) above, the fundamental issue of EGT is that the beliefs that the players were supposed to use at step (i) are defined from the choice observed by the theorist at step (iii). Indeed, once we identified an equilibrium, we can build the *ad hoc* prior beliefs such that Bayes rational players should play this equilibrium. This however does not give any practical advice to a player. EGT only suggests that, if Bayes rationality is common belief, then the possible sets of prior beliefs the players have at their disposal is reduced (and corresponds to correlated equilibrium distributions if we also assume common priors for instance). Endorsing this approach makes deliberation 'vacuous' (Levi, 1998, p. 181). Indeed, deducing my beliefs from the equilibrium profile of strategies implies that the premises of my deliberation (my prior beliefs) is deduced from its result (my actual choice).

Since the behaviouristic interpretation of payoffs provides an inconsistent account of beliefs, we will endorse a mentalistic account of payoffs, which requires introducing a complementary theory of belief formation. We suggest integrating the ST in the model to explain how players form their beliefs in step (i): we will then assume that players maximise their expected payoff in step (ii), given their beliefs in step (i), and characterise the resulting solution concept as a *subjective belief equilibrium*. Unlike the behaviouristic approach that deduces solution concepts from vNM utilities (i.e. from the choice itself), deriving a solution concept from a mentalistic notion of 'material payoff' is not tautological (Lehtinen, 2011, pp. 277-278).

2. Simulation theory and the massaging process

2.1 Simulation and the formation of beliefs

Predicting someone else's behaviour requires forming expectations about the mental states leading this other to adopt a specific behaviour. In a game situation, this typically means forming expectations about her preferences, beliefs, and objectives. Once P1 has formed her beliefs about the choice of P2 (based on what P1 thinks P2's preferences, beliefs, and objectives are), P1 can determine her best reply. The simulation theory suggests that P1 uses her own mind to predict the behaviour of P2: 'our own mental processes are treated as a manipulable model of other minds' (Goldman and Shanton, 2012, p. 10). Simulation is therefore an efficient heuristic for predicting someone else's decision (Shanton and Goldman, 2010) – a recent upsurge of empirical data provided by neuro-imaging indeed contributes to suggest that simulation is a very effective process of mindreading (Goldman, 2006; see also Singer and Fehr, 2005; Kirman and Teschl, 2010). Consider the Stag Hunt discussed above (with material payoffs rather than vNM utils):

SH	A_2	B_2
A_1	(\$3;\$3)	(\$0;\$2)
B_1	(\$2;\$0)	(\$1;\$1)

Suppose that P1 intends to play a best reply to P2's strategy: P1 must therefore anticipate the choice of P2 (strategy A_2 or strategy B_2). To do so, she imagines herself in P2's shoes. The simulation process is then a three steps process (P1 is called the *attributor* or *simulator*, and P2 the *target*). First, P1 brings forth 'pretend' or 'imaginary' mental states – like the intention for P2 to maximise her material payoff, the preference for the profile A_1A_2, the belief that P1 is likely to choose A_1 – in her own mind. Those mental states are supposed to 'mimic' those of her target: as a simulator, P1 pretends that those imaginary intentions, preferences and beliefs, are those of her target P2.[v] Second, she feeds these pretend states in her own decision-making system, which runs 'off-line'. As a best-reply reasoner, P1 imagines what she would play if she were choosing instead of P2, with the intention of reaching a payoff maximising profile for P2, given P2's belief that P1 is likely to choose (say)[vi] A_1. The output of the simulation process (the best reply to what P1 thinks P2 believes about P1) is therefore A_2. Lastly, this output is attributed to her target (Goldman, 2006, p. 20). P1 believes that P2 will choose A_2, her best reply to P2's choice is thus to play A_1. Simulation thus explains how P1 forms her beliefs about P2's mental states, reasoning, and choice.

If the attributor's decision-making process and pretend initial mental states are similar to that of the target, then the output of the simulation process is a reliable prediction of the target's choice. Simulation may however not lead to accurate predictions, for instance if the mental states attributed to the other are incorrect, and 'chosen badly out of ignorance' (Goldman, 2006, p. 48), or if the other's decision process is different from the one of the simulator. Numerous experimental findings indeed report *egocentric biases* in predicting other's choices, which preclude from an accurate mindreading (Goldman, 2006, pp. 177-179). In the previous illustration, although P1 predicts that P2 will choose A_2, it is not certain that P2's actual beliefs about the choice of P1 correspond to the 'pretended' beliefs simulated by P1. Furthermore, unlike P1 who is a best reply reasoner, P2 could be a maximin player, i.e. a player who always chooses her maximin strategy (B_2 in the Stag Hunt). Since P1 and P2 reasoning processes are different, it is likely that the outcome of their reasoning will be different. In cases of erroneous predictions, the simulators revise their beliefs about their targets' mental states, and then run other simulations with different inputs.

In any case, even with very little information about the target, such as the information given in a payoff matrix, an individual can form a prediction, based on her own perception of the game, and her own cognitive scheme and reasoning process. Hence, the ST provides a theory to derive endogenously the beliefs of the players from the structure of the game.

2.2 Reaching consistent priors: the massaging process

Note that the specificity of game theory – which was put forward by Harsanyi and then Aumann against Kadane and Larkey – is that the players expect each other to act rationally. This means that the beliefs of P1 are already the outcome of a reasoning process, because they must be revised to be consistent with the rationality of P2. In Bacharach and Hurley's (1991, p. 26) words, the priors are thus 'the central unknowns of the theory':

> What brings me to have the prior probabilities that I do for your deciding on one option and another is a question not answered (and rarely asked) by the Bayesian theory of games. The absence of an independent account of what is in the players' priors is a grave lacuna. There are many games for which, once the priors are given, the identities of the rational acts follow trivially, and then game theory itself is trivialized if it is merely assumed that the prior are such and such. To avoid this trivialization by Bayesianization, we must take the content of the priors in such cases to be the central unknowns of the theory, endogenous to it.

We will use here Binmore's (2009, pp. 130-132) description of the 'massaging process' to model how P1 could reach consistent[vii] priors. Binmore relies on Savage's distinction between small and large worlds to question Bayesianism – which he defines as 'the philosophical position that Bayesian methods always applies to all decision problems' (Binmore, 2009, p. 96) – and in particular the claim that 'rationality endows agents with prior probabilities' (Binmore, 2006, p. 3). Binmore on the contrary intends to explain the formation of prior beliefs, while remaining faithful to Savage's theory. Rather than directly using my 'gut feelings' (Binmore, 2009, p. 130) at the prior stage to form my beliefs, I should imagine what would be my gut feelings after the realisation of the state of nature. For each realisation, I deduce a posterior belief from my hypothetical gut feelings: it is however unlikely that those posterior beliefs are consistent. I should then 'massage' my posteriors until I reach consistent posteriors (i.e. revise my posteriors, knowing that my hypothetical gut feelings were inconsistent). Once I reached consistent *massaged* posteriors, Bayes' rule guarantees that they can be deduced from a prior. The formation of my prior belief thus requires a stage of introspection and self-reflection, and this is precisely what we will do with ST, by explicitly modelling the introduction of the common belief in rationality in the players' beliefs *via* a similar 'massaging process'.

Binmore quickly mentions how the massaging process could work in game theory. Rather interestingly, he suggests that 'Alice will then not only have to massage her own probabilities until consistency is achieved, she will also have to *simulate* Bob's similar massaging efforts' (Binmore, 2009, p. 135, our emphasis), and suggests that the resulting beliefs form a subjective equilibrium, which should necessarily be a Nash equilibrium. Although our proposition 4 will confirm Binmore's intuition, we will also show that the introduction of simulation (which is simply a passing remark in Binmore's argument, and not a well-developed game theoretical analysis) could justify a much larger set of equilibria (that we will call 'subjective belief equilibrium' instead of 'subjective equilibrium'). The reason is that, even when massaging our priors until they become consistent, it is possible to form *action-dependent beliefs* (that Binmore explicitly rejects for reasons we will discuss in section 4).

As an illustration of the formation of action-dependent beliefs through the massaging process, consider the previous Stag Hunt:

SH	A_2	B_2
A_1	(\$3;\$3)	(\$0;\$2)

B_1	(\$2;\$0)	(\$1;\$1)

Unlike EGT – according to which P1 is endowed with prior probabilities, which should necessarily be consistent with the mutual rationality of the players – we suppose that P1 imagines what would be her gut feelings about the action of P2 after she chose her own action. In the spirit of Haruvy, Stahl and Wilson (1999), suppose that each player can either be 'optimistic' or 'pessimistic'. An optimistic type 'tends to choose the strategy which can potentially give him the highest payoff for a given game' (p. 256) and a pessimistic type 'is motivated by worst case scenarios and hence tends to choose a secure action' (p. 257). An optimist would therefore select the payoff dominant equilibrium, while a pessimist would opt for a maximin strategy, and therefore the risk dominant equilibrium. P1 can thus be motivated by two conflicting reasons: either (R1) try to reach the highest payoff (maximax) or (R2) secure the highest minimum payoff (maximin). P1 however does not necessarily know what psychological factors make her privilege a reason over another.[viii] Therefore, if P1 eventually chooses A_1, she may deduce that the reason (R1) appeared as relatively more important than (R2) for some unknown psychological reason. When simulating the reasoning of P2, she will assume that the same psychological forces are driving P2's choice (since she attributes her own mental states to P2 when simulating her reasoning). From that prospect, if P1 chooses A_1, she may attribute a higher probability for P2 choosing A_2, because she will assume that the psychological forces that pushed her to privilege (R1) and then to choose A_1 are also probably operating in P2's mind. P1 can then form an action-dependent belief, not because she thinks her choice *directly* influences the choice of the other, but because she is aware that her choice is driven by psychological factors that could also influence P2. Her gut feelings about the choice of the other could indeed depend on her own actions: if

P1 feels optimistic, and thus chooses A_1, she will more likely believe that P2 chooses A_2, because she tends to attribute her own optimism to P2. On the contrary, when considering what would be her gut feelings about P2's choice if she chose B_1, P1 will attribute her own pessimism to P2, and is thus more likely to believe that P2 will choose B_2. P1's prior belief could therefore be of the following form:

SH	A_2	B_2
A_1	α	0
B_1	0	$1-\alpha$

In this case, P1 believes that the outcome of the game is A_1A_2 or B_1B_2, with probability α and $(1-\alpha)$ respectively. Those priors simply state that both players are best reply reasoners (i.e. P1 plays A_1 if and only if P2 plays A_2). We thus obtain correlated priors, although they are consistent with the mutual rationality of the players – we even have in this example a correlated equilibrium distribution.

The massaging process can therefore offer a *psychological* explanation of the formation of consistent (and possibly correlated) prior beliefs. The player indeed refers to her own psychological make-up (her gut feelings) to form her prior beliefs, which she revises until they are consistent with the rationality of the other players.

3. Subjective belief equilibrium

We now present the concept of *subjective belief equilibrium*, as a strategy profile that results from the joint maximisation of individual expected payoffs, when the beliefs of the players have been 'massaged' by the players. We start by introducing ST in a game-theoretical framework (section 3.1), and then characterise a *massaged belief hierarchy* as the beliefs formed by Bayes rational players who simulate the reasoning of the other players (section 3.2). We then illustrate the possibility of forming consistent action-dependent beliefs (section 3.3) and define a subjective belief equilibrium as the payoff maximising profile derived from massaged beliefs (section 3.4).

3.1. Simulation theory in games

We introduce the following belief operators:

- 1st order belief: $B_i(E)$ means 'i believes *E*', with $E \in \mathcal{E}$ a proposition, and $\bar{E}$ its negation
- 2nd order belief: $B_{ij}(E) = B_i\left(B_j(E)\right)$, means '*i* believes that *j* believes *E*'
- Mutual belief: $MB(E) = B_i(E) \cap B_j(E) = (B_i \cap B_j)(E)$ means that 'i and j believes E'
- Common belief: $CB(E) = (\cap_{k=1}^{\infty} MB^k)(E)$ means that 'E is mutual belief', that 'the proposition "E is mutual belief" is mutual belief', and so on *ad infinitum*
- Uncertainty: $U_{ij}(E) = \overline{B_i\left(B_j(E)\right)} \cap \overline{B_i\left(\overline{B_j(E)}\right)}$ means that *i* neither believes that '*j* believes *E*', nor that '*j* does not believe *E*'.

We saw that the simulation routine (i) tends to imply an egocentric bias, i.e. that the individual tends to attribute her *own* beliefs and perceptions to the other, and (ii) that the simulator uses her own reasoning process to simulate the reasoning of her target. An egocentric bias means that, when *i* is in a situation of uncertainty regarding *j*'s knowledge of a proposition *E* (or at least, *i* does not believe that *j* does not believe *E*), then *i* tends to assume that j believes E. Formally, we can translate this property as follows:

$$\text{SIMB}_\text{i}: \quad B_i(E) \cap U_{ij}(E) \Longrightarrow B_{ij}(E). \tag{1}$$

SIMB_i means that, in presence of uncertainty, *i* attributes her own beliefs to player j. Think for instance of a coordination game with a focal point: if I believe that an option is more salient than the others, and that I don't have a specific reason to believe that you don't perceive this option as more salient, then I will assume you also perceive this outcome as the most salient (which could then rationalise our coordination on the focal point).

When the uncertainty extends to the beliefs of another player, i.e. $U_{ij}(E'), \forall E' \in \mathcal{B}_j$ with $\mathcal{B}_j$ the set of *j*'s beliefs about the beliefs of other players (i.e. *j*'s first-order beliefs, second-order beliefs, etc.), we have the following result – similar to Friedell's (1969, p. 31) intuition of the emergence of *common opinion* – (all the proofs are given in appendix):

Proposition 1. Let E be a proposition, and suppose that $U_{ij}(E'), \forall E' \in \mathcal{B}_j$:

$$B_i(E) \cap U_{ij}(E) \cap SIMB_i \Longrightarrow B_i\big(CB(E)\big). \tag{2}$$

Proposition 1 means that, if *i* believes *E* and is uncertain about *j*'s beliefs (i.e. about *j*'s belief about *E*, but also about *j*'s belief about *i*'s belief about *E*, etc.), then *i* believes that *E* is common belief among them. Unless *i* has a good reason to believe that *j* does not believe *E*, *i* simply assumes that '*j* is in the same cognitive position as *i* himself' (Friedell, 1969, p. 31): *i* will therefore believe that they both believe *E*, but also that they both believe that they believe *E*, etc.. Since *i* is uncertain about *j*'s beliefs, and in particular if *j* believes that they both believe *E*, *i* indeed also attributes her second-order belief to *j*, etc.. Friedell (1969) then shows that this 'symmetry' of cognitive positions between the players generates a structure of common belief.[ix]

A corollary of proposition 1 is that, if all the players attribute their own beliefs to the others in presence of uncertainty, and if a proposition E is mutual belief among the players, then it is also common belief among them (by proposition 1, they indeed all believe that E is common belief, i.e. that the common belief of E is mutual belief, which is the same thing than the common belief of E). An interesting implication of this corollary for game theory is that mutual knowledge of the structure of the game is sufficient to ensure its common belief. Although the assumption of common knowledge of the structure of the game seems too cognitively demanding, simulation and mutual knowledge of the structure of the game are sufficient to generate a formally equivalent epistemic structure. The insight of simulation is indeed that we do not require the *actual* belief of the players, but simply that they have the cognitive capacity to *generate* it.

Regarding the simulation of the reasoning process of the other individual, we must firstly introduce player *i*'s *choice function*. Consider a choice problem $\mathrm{P} = (A_i, C_i, O_i)$, in which *i* must choose an action $a_i \in A_i$, with a consequence $C_i(a_i)$, to satisfy her objective O_i (e.g. maximise her monetary gain). We define a choice function $\mathcal{C}_i: \mathrm{P} \longmapsto A_i$ as a function that associates to a choice problem $\mathrm{P} = (A_i, C_i, O_i)$ the action to be chosen in A_i. Simulation means that, for a given choice problem P_j for player *j*:

$$\mathrm{SIMR_i}: \quad \mathcal{C}_i(\mathrm{P}_j) = f(A_j, C_j, O_j) \Rightarrow B_i\left(\mathcal{C}_j(\mathrm{P}_j) = f(A_j, C_j, O_j)\right). \tag{3}$$

$\mathrm{SIMR_i}$ means that, if the function $f: A_j \times C_j \times O_j \longmapsto A_j$ corresponds to the function player i would apply to select an action in A_j (if she were in the position of player *j*, i.e. if she had to choose instead of *j*, given *j*'s objectives), then *i* believes that *j* would apply the same function to select an action (e.g. if my objective is to maximise my monetary gain, then f would be the argmax function applied to *j*'s monetary gain).

3.2 Massaged belief hierarchy

We suggested in section 2.2 that simulation could lead players to form action-dependent beliefs (ADB). We now show that it is also possible for the players to *rationalise* those beliefs, i.e. to hold consistent ADB. We introduce the notion of *massaged belief hierarchy* as the beliefs of the players resulting from the simulation of each other's reasoning.

We consider finite games in normal forms, i.e. games $G = \langle N, X, \Pi \rangle$, with $N = \{1; \dots; n\}$ the set of players, $X = \prod_{i \in N} X_i$ with X_i the finite set of player *i*'s strategies, and $\Pi_i: X \longmapsto \mathbb{R}$ player *i*'s material payoff. As discussed above, this material payoff function is the primitive of the game, and we assume that maximising one's material payoff is the objective the individual intends to achieve (in a non-tautological way). $\Delta(X) = \left\{\{P(x)\}_{x \in X} \in [0;1]^{|X|} \middle| \sum_{x \in X} P(x) = 1\right\}$ denotes the set of discrete probability distributions over *X*. While *x* denotes players' actions, we use the letter *s* to denote players' beliefs about actions. $s_{i,j,k} \in X_k$ denotes for instance *i*'s belief about *j*'s belief about *k*'s strategy:

$$s_{i,j,k} = \bar{x}_k \iff B_{ij}(x_k = \bar{x}_k). \tag{4}$$

For a given profile of conditional probability distributions $P(X) = \{P(X_i|X_{-i})\}_{i \in N}$, we denote by $\Omega(\{P(X_i|X_{-i})\}_{i \in N})$ the set of probability distributions that can represent $\{P(X_i|X_{-i})\}_{i \in N}$. $\Phi_{\mathrm{i}}(X)$ denotes the set of conditional probability distribution $P(X_i|X_{-i})$ for player *i*. As will be illustrated in section 3.3, the function $\Omega: \Phi \longmapsto \Delta(X)$ is neither injective nor surjective: a single set of conditional probability distributions can indeed be represented by several probability distributions over outcomes, and several sets of conditional probability distributions can be represented by the same probability distribution. Lastly, two conditional probabilities $P(X_i|X_{-i})$ and $P(X_{-i}|X_i)$ are not necessarily compatible, i.e. there may not exist a distribution $p \in \Delta(X)$

that can simultaneously represent those two conditional probabilities (see e.g. Arnold and Press, 1989) – in which case $\Omega(\{P(X_i|X_{-i})\}_{i\in N}) = \emptyset$.

We denote by $S_i = \left\{\left\{s_{i,j}\right\}_{j\neq i}; \left\{s_{i,j,k}\right\}_{k\neq j;\, j\neq i}; \left\{s_{i,j,k,l}\right\}_{l\neq k;\, k\neq j;\, j\neq i}; \ldots\right\}$ player i's belief hierarchy, i.e. the infinite set including player i's 1st-order beliefs $s_{i,j}$ (her belief about *j*'s strategy), her 2nd order beliefs $s_{i,j,k}$ (her belief about *j*'s belief about *k*'s strategy), etc. Our objective is to identify the belief hierarchies that a Bayes rational player could hold if she simulates the reasoning of the other players (we suppose throughout the rest of the section that players are initially uncertain about the beliefs of the others, and must in consequence form those beliefs by simulating their reasoning). Suppose that players $i \in N$ choose the strategy that maximises their expected material payoff:

$$\text{PM: } \max_{x_i\in X_i}\left[\sum_{x_{-i}\in X_{-i}} s_{i,-i}\,(x_{-i}|x_i)\,\Pi_i(x_i,x_{-i})\right] \qquad (5)$$

with $s_{i,-i}\colon X_i \longmapsto \Delta(X_{-i})$ *i*'s belief about the strategy of the players in -*i*.

Definition. S_i is a massaged belief hierarchy if and only if PM, $SIMB_i$ and $SIMR_i$ are true for player i.

A massaged belief hierarchy (MBH) is the belief hierarchy of player *i* when she simulates the reasoning of other players. Those beliefs must be compatible with the rationality of the other players, since *i* will assume that the others are expected payoff maximisers if she is herself an expected payoff maximiser (which is true by PM). We can now characterise a MBH:

Proposition 2. *S_i is a massaged belief hierarchy if and only if:*

(i) $s^*_{i,[k],j} = s^*_{i,j}$ with $j \neq i$, for all sequences of players $[k] = k_1, k_2, \dots, k_m$.

(ii) $s^*_{i,[k],i} = s^*_{i,j,i}$, $\forall j \neq i$, for all sequences $[k] = k_1, k_2, \dots, k_m$.

(iii) *and there exists $s^* \in \Omega\left(\{s^*_{i,j,k}\}_{i\neq j; j\neq k}\right)$ such that:*

$$\sum_{x\in X} s^*(x)\Pi_k(x) \geq \sum_{x\in X} s'(x)\Pi_k(x), \qquad \forall s' \in \Omega\left(\{s^*_{i,j,k}\}_{i\neq j;j\neq k}; s'_{i,j,k}\right), \forall s'_{i,j,k} \in \Phi_k, \forall k \in N. \quad (6)$$

Condition (i) means that *i* believes that her belief about *j*'s strategy is shared by all the other players, and that it is common belief (the *m*th order belief of *i* about *j*'s strategy is indeed always equal to her first order belief about *j*'s strategy). Condition (ii) is similar to (i), since it means that *i* believes that all the players have the same belief about her own strategy, and that it is common belief. Those two conditions ensure that the beliefs of the players converge, in the sense that *i* believes that the first order beliefs of all the players are common belief and identical. It also implies that it is sufficient to work with *i*'s 1st and 2nd order beliefs, rather than with the whole belief hierarchy S_i.

Condition (iii) means that the expected payoff of player *k* at the MBH (i.e. if the choice of the players is accurately described by the distribution s^*) should be at least equal to her expected payoff if she 'deviates' to another conditional distribution $P'(X_k|X_{-k})$. In other words, a MBH is the probability distribution $P^*(X)$ induced by a set of conditional probability distributions $\{P^*(X_i|X_{-i})\}_{i\in N}$, such that $P^*(X_i|X_{-i})$ is maximising the expected utility of player *i*, $\forall i \in N$.[x]

The intuition supporting condition (iii) is the following. Consider *i*'s belief about the strategy of player *j*:

$$s_{i,j}\left(x_{-j}\right) = P\left(X_j \middle| X_{-j} = x_{-j}\right) \tag{7}$$

This belief is rationalisable if and only if *i* can justify why *j*'s choice is accurately described by $P\left(X_j \middle| X_{-j} = x_{-j}\right)$. This means that $s_{i,j}$ must be a 'rational' choice for player *j*, given *j*'s belief about the choice of players in – *j*. Given S_i, *i*'s belief about *j*'s belief about the strategy of players $k \in -j$ is:

$$s_{i,j,-j}\left(x_j\right) = P\left(X_{-j} \middle| X_j = x_j\right) \tag{8}$$

If there exists a distribution $P'\left(X_j \middle| X_{-j}\right) \neq P\left(X_j \middle| X_{-j}\right)$ such that:

$$\sum_{x \in X} s(x)\Pi_j(x) > \sum_{x \in X} s'(x)\Pi_j(x), \tag{9}$$

with $s' \in \Omega\ \left(P'\left(X_j \middle| X_{-j}\right); P\left(X_{-j} \middle| X_j\right)\right)$, then it means that $s_{i,j}$ is not a rationalisable belief from *i*'s perspective, because *j* would be better off if her choice was described by $P'\left(X_j \middle| X_{-j}\right)$. The belief hierarchy of *i* should therefore be such that she can rationalise the beliefs she attributes to the other players (this is how the player 'massages' her prior so as to reach consistent posteriors): each conditional distribution $P\left(X_j \middle| X_{-j}\right)$ should therefore maximise the expected utility of player *j*, given *i*'s belief about *j*'s belief about the strategy of the players $k \in -j$ (and these higher order beliefs must also be rationalisable – this is ensured by parts (i) and (ii) of the proposition, because they are the same than $s_{i,j}$ and $s_{i,j,i}$).

We can show the following existence result for a MBH:

Proposition 3. *Let $G = \langle N, X, \Pi \rangle$ be a game in normal form. If G has a Nash equilibrium $p^* \in \Delta(X)$, then there exists a massaged belief hierarchy S_i^* for each player i.*

Proposition 3 ensures that the existence of a Nash equilibrium is a sufficient condition for the existence of a massaged belief hierarchy for each player. In other words, if a Nash equilibrium exists, then we know that the players can always manage to rationalise a belief hierarchy S_i^*.

We want to emphasise that the optimisation condition defining a MBH is about players' *beliefs*, and not their actual choices. Condition (6) indeed suggests that player *i*'s strategy can be conditional on player *j*'s strategy (since we look for a *conditional* probability distribution that maximises *i*'s expected material payoff): this is only because player *j* believes that her choices and those of player *i* could be correlated (and *j* can believe this if she believes that *i* believes it). *When forming their beliefs* (and not when choosing their actual strategies), players therefore test the consistency of their beliefs not with regard to their own actions (which are independent), but with regard to their beliefs about the others (which may be dependent). Although all the players know that they cannot choose conditional distributions, they can use the fact that all the players can believe that beliefs are action dependent, and then form a belief hierarchy in which beliefs are correlated. Furthermore, since this massaging process only happens in *i*'s mind, nothing guarantees that all the players will reach the *same* MBH (in particular when several distributions satisfy condition iii). A MBH is therefore an individual concept, and does not guarantee that the players' actual choice will be accurately described by the MBH (the methodological implications of the multiplicity of MBH will be discussed in the conclusion).

3.3 Illustration: Prisoner's dilemma

As an illustration, consider the following Prisoner's dilemma (the payoffs in the matrix should be interpreted as material payoffs and not as vNM utilities):

PD	C_2	D_2
C_1	(2;2)	(0;3)
D_1	(3;0)	(1;1)

Consider the following belief hierarchies for player 1 (we only focus on 1st and 2nd order beliefs, and assume that higher order beliefs will be identical):

- $\boldsymbol{S_1^1}$: $s_{1,2}(C_2|X_1) = 0$, $s_{1,2,1}(C_1|X_2) = 0$, $\forall X_i \in \{C_i, D_i\}$. P1 believes that P2 never cooperates, and believes that P2 believes that P1 never cooperates,
- $\boldsymbol{S_1^2}$: $s_{1,2}(C_2|X_1) = 1$, $s_{1,2,1}(C_1|X_2) = 1$, $\forall X_i \in \{C_i, D_i\}$. P1 believes that P2 always cooperates, and believes that P2 believes that P1 always cooperates,
- $\boldsymbol{S_1^3}$: $s_{1,2}(C_2|C_1) = 1$, $s_{1,2}(C_2|D_1) = 0$, $s_{1,2,1}(C_1|C_2) = 1$, $s_{1,2,1}(C_1|D_2) = 0$. P1 believes that P2 is ready to cooperate if and only if P1 cooperates, and believes that P2 believes that P1 is ready to cooperate if and only if P2 cooperates (P1 believes that P2 is a conditional cooperator, and believes that P2 believes that she is also a conditional cooperator),
- $\boldsymbol{S_1^4}$: $s_{1,2}(C_2|C_1) = 0$, $s_{1,2}(C_2|D_1) = 0$, $s_{1,2,1}(C_1|C_2) = 1$, $s_{1,2,1}(C_1|D_2) = 0$. P1 believes that P2 always defect, and believes that P2 believes that P1 is a conditional cooperator,
- $\boldsymbol{S_1^5}$: $s_{1,2}(C_2|C_1) = 0$, $s_{1,2}(C_2|D_1) = 1$, $s_{1,2,1}(C_1|C_2) = 1$, $s_{1,2,1}(C_1|D_2) = 0$. P1 believes that P2 cooperates if and only if P1 defects, and believes that P2 believes that P1 is a conditional cooperator.

We now represent each belief hierarchy S_1 by a probability distribution $s \in \Omega(s_{1,2}; s_{1,2,1})$ and check whether they form a MBH.

$\boldsymbol{S_1^1}$	C_2	D_2
C_1	0	0
D_1	0	1

Given that P1 believes that P2 always defects and believes that P2 believes that P1 always defects, the only possible representation s of those conditional distributions is that the probability of reaching D_1D_2 is equal to 1. Given those beliefs, the expected payoff for both players (from P1's perspective, since we only consider P1's belief hierarchy) is 1. P1 knows that, if P2 believes that P1 always defects, then P2 can only get a payoff of 0 (if she cooperates) or 1 (if she defects). This means that any other conditional probability distribution than $s_{1,2}(C_2|X_1) = 0$ cannot give a strictly higher payoff to P2. $s_{1,2}$ is therefore a rationalisable belief, since believing $s_{1,2,1}(C_1|X_2) = 0$ implies that $s_{1,2}(C_2|X_1) = 0$ is payoff maximising (if P1 believes $s_{1,2,1}(C_1|X_2) = 0$, then she can explain why P2 would unconditionally defects). By a similar argument (since we assumed that $s_{1,2,1,2} = s_{1,2}$) we could rationalise P1's 2$^{\text{nd}}$-order belief, i.e. that P2 believes that she always defects. This is indeed payoff maximising, given P1's belief about P2's beliefs.

Since P1 can rationalise the beliefs about P2's strategy and the beliefs she attributes to P2, S_1^1 is a massaged belief hierarchy (if players could choose conditional strategies instead of pure strategies, then no player could strictly increase her material payoff by switching to another conditional strategy).

$\boldsymbol{S_1^2}$	C_2	D_2
C_1	1	0
D_1	0	0

We now consider the second belief hierarchy, according to which both players always cooperate. If P1 believes that P2 believes that P1 always cooperates, then P1 cannot also believes that P2 always cooperates: if P2 believes that P1 always cooperates, then it would be in the interest of P2 to always defect. P1 must therefore revise her inconsistent beliefs: $\boldsymbol{S_1^2}$ is not a massaged belief hierarchy.

$\boldsymbol{S_1^3}$	C_2	D_2
C_1	α	0
D_1	0	$1-\alpha$

Unlike the two previous belief hierarchies, $\boldsymbol{S_1^3}$ can be represented by several distributions $s \in \Omega(s_{1,2}; s_{1,2,1})$. Since P1 believes that it is common belief that P1 and P2 are conditional cooperators, the two only possible outcomes of the game are mutual cooperation and mutual defection. There is however not enough information to determine the *frequency* of each outcome: we have therefore $\alpha \in [0; 1]$. We can check that the expected payoffs for both players is $\big(2\alpha + (1-\alpha)\big) = 1 + \alpha$.

Since P1 believes that it is common belief that P1 and P2 defects if the other defects, we can check that P1 cannot believe that P2 would always choose to defect rather than conditionally cooperate (since P1 would then defect, and P2 would only get a payoff of 1). Note however that, if $\alpha < 1$ (players do not necessarily coordinate on the cooperative outcome), P1 cannot believe that P2 is a conditional cooperator: it would indeed be in P2's interest to always cooperate (guaranteeing a payoff of $2 > (1+\alpha)$) –

P1 would therefore revise her beliefs. Nevertheless, if $\alpha = 1$ P1 knows that being a conditional cooperator is payoff maximising for P2 (even though being an unconditional cooperator would also be payoff maximising – but in this case P1 could not simultaneously believe that P2 always cooperates, and that P2 believes that P1 is a conditional cooperator, because P2 should realise that being a conditional cooperator is not payoff maximising for P1).

$\boldsymbol{S_1^2}$ is therefore a massaged belief hierarchy, because there exists a distribution $s \in \Omega(s_{1,2}; s_{1,2,1})$ such that P1 can rationalise the common belief that both players are conditional cooperators (when $\alpha = 1$). It is noticeable that, although the representations of the belief hierarchy $\boldsymbol{S_1^2}$ and $\boldsymbol{S_1^3}$ are the same (the profile $(C_1; C_2)$ occurs with probability 1), only S_1^3 is a MBH: it is indeed the conditional distributions supporting the distribution over strategy profiles (and not the distribution itself) that are relevant for characterising a MBH – mutual cooperation is rationalisable only if both players are conditional cooperators.

$\boldsymbol{S_1^4}$	C_2	D_2
C_1	0	0
D_1	0	1

In the present case, although P1 believes that P2 believes that P1 is a conditional cooperator, P1 believes that P2 always defect. The only possible outcome is therefore (D_1, D_2), with an expected payoff of 1 for both players. However, P1 cannot simultaneously believe that P2 (i) always defect and (ii) believes that P1 is a conditional cooperator. If P2 believes that P1 is a conditional cooperator, then P2 would be better off by always cooperating (the probability of (C_1, C_2) would be 1), or becoming a conditional cooperator (with a probability of (C_1, C_2) of α). $\boldsymbol{S_1^4}$ is not a massaged belief hierarchy, because P1 cannot rationalise her beliefs $s_{1,2}$.

The last belief hierarchy $\boldsymbol{S_1^5}$ is a bit peculiar, because it suggests that one player could intend to cooperate if and only if the other defects, i.e. that the objective of this player is to reach only asymmetric payoffs. The main issue is however that this 'asymmetric' behaviour and being a conditional cooperator are *incompatible* probability distributions. It just means that it is not possible to represent in a matrix a probability distribution consistent with both conditional probability distributions. Since $\Omega(s_{1,2}, s_{1,2,1}) = \emptyset$, $\boldsymbol{S_1^5}$ is not a massaged belief hierarchy.

3.4 Subjective belief equilibrium

We now define a solution concept representing the choice of Bayes rational players who massaged their beliefs. We define a *subjective belief equilibrium* as the strategy profile resulting from the maximisation of the players' subjective expected payoff. This equilibrium is restricted to *pure* strategies, because – in line with Aumann (1987, p. 15) – we interpret mixed strategy equilibria in terms of beliefs. The players cannot randomize when choosing their strategy, but their belief hierarchy can include mixed strategies if they are not certain of the actions of the other players. In our model, a mixed strategy Nash equilibrium would be the (common) massaged belief hierarchy of the players, while the subjective belief equilibrium would correspond to the strategy profile resulting from the maximisation of their expected payoff (if they accurately anticipate it).

Definition. A strategy profile $x^* \in X$ is a subjective belief equilibrium if and only if, $\forall i \in N$, there exists a massaged belief hierarchy S_i^* such that:

(i) $\sum_{x_{-i} \in X_{-i}} s^*(x_{-i}|x_i^*)\ \Pi_i(x_i^*; x_{-i}) \geq \sum_{x_{-i} \in X_{-i}} s^*(x_{-i}|x_i')\ \Pi_i(x_i'; x_{-i}),\ \forall x_i' \in X_i$

(ii) $s^*(x_{-i}|x_i^*) = x_{-i}^*,\ \ \forall i \in N, i \neq j.$

With $s^* \in \Omega\left(\left\{s_{i,j,k}^*\right\}_{i \neq j; j \neq k}\right)$ the representation of S_i^*given by proposition 2.

The first condition means that each player chooses the strategy that maximises her expected payoff, given her beliefs about the action of the others. Those beliefs should however be rationalisable, since they are derived from a massaged belief hierarchy. The second condition means that the players accurately predict the choice of the other players. This means that the existence of a MBH is a necessary but not sufficient condition for the existence of a SBE. For instance, in the 'matching pennies' game, the two possible strategies have the same expected payoff (if players' beliefs are such that they believe it is common belief that they play each strategy with probability ½, which is a MBH), but the players cannot predict the actual choice of the other player.

Similarly to Binmore's (2009, p. 135) suggestion, we can establish a direct connection between Nash and a subjective belief equilibrium:

***Proposition 4**. A Nash equilibrium in pure strategies is a subjective belief equilibrium.*

The intuition of the proof is quite straightforward and follows from proposition 3: if there exists a Nash equilibrium in pure strategies, then the players can form the consistent priors that they all play this equilibrium. Given this prior belief, they maximise their expected payoff by playing their part of the Nash equilibrium. Unlike Binmore, we will however show that the set of subjective belief equilibria is much larger, since it is possible to form consistent ADB.

4. Simulation, ratifiability, and action-dependent beliefs

We now highlight that the introduction of ST will not only provide a psychological explanation of belief formation, but could also undermine the implicit but problematic identification of Bayes rationality with best-reply reasoning in game theory.

The most common requirement of epistemic game theory – its 'central idea' according to Perea (2014, p. 13) – is the *common belief in rationality* (CBR), i.e. that it is common belief that players choose the strategy that maximises their expected utility. A widely accepted proposition is that CBR implies the iterated deletion of dominated strategies (see e.g. Bernheim, 1984), meaning that a rational player cannot believe that another rational player could choose an iteratively dominated strategy. This result however requires the additional (and often implicit)[xi] assumption of *ratifiability* (Jeffrey, 1990; Levi, 1998), according to which the actions of the players $j \neq i$ are probabilistically independent of *i*'s choices. Formally:

Definition. Let $p \in \Delta(X)$ be a probability distribution over X. A strategy profile $x^* \in X$ is ratifiable *if and only if*, $\forall i \in N$*:*

$$\sum_{x_{-i} \in X_{-i}} p(x_{-i}|x_i^*)\ \Pi_i(x_i^*; x_{-i}) \geq \sum_{x_{-i} \in X_{-i}} p(x_{-i}|x_i^*)\ \Pi_i(x_i'; x_{-i}), \qquad \forall x_i' \in X_i. \quad (10)$$

A strategy is ratifiable if and only if it gives a higher expected payoff than any other strategy $x_i' \in X_i$, while the probability $p(x_{-i}|x_i^*)$ defining the expected payoff remains the same after the deviation to x_i'. Jeffrey (1990) however argues that ratifiability is not implied by Bayesian rationality, since the maximisation of one's expected payoff implies:

$$\sum_{x_{-i} \in X_{-i}} p(x_{-i}|x_i^*)\ \Pi_i(x_i^*; x_{-i}) \geq \sum_{x_{-i} \in X_{-i}} p(x_{-i}|x_i')\ \Pi_i(x_i'; x_{-i}), \qquad \forall x_i' \in X_i. \quad (11)$$

When deviating to x_i', player *i* must consider the possible impact of her choice on the state of the world, i.e. on the strategies of the other players (the posterior distribution $p(x_{-i}|x_i')$ is therefore not necessarily equal to $p(x_{-i}|x_i^*)$ – see Mariotti, 1996, pp. 143-144, for a similar point).

A reason why ratifiability is often conflated with Bayesian rationality is that believing that the actions of the others could depend on one's own action – and therefore that $p(x_{-i}|x_i^*) \neq p(x_{-i}|x_i')$ – seems to be a fallacious mode of reasoning. If Bayesian rationality is common belief among us, we should know that our decisions are independent, and therefore that our choices cannot directly influence the choices of others (e.g. Binmore, 1992, pp. 311-312). The direct implication is that players can never play strictly dominated strategies, such as cooperating in a prisoner's dilemma: whatever my belief is about your strategy, defecting is always payoff maximising. The common belief of Bayes rationality and ratifiability therefore implies the elimination of iteratively dominated strategies.

However, since our objective is to develop a *psychological* theory of belief formation in games, we should also consider the possibility that players believe that their actions are correlated,[xii] and accordingly that their beliefs about the action of others may depend on their *own* actions. ADB could indeed explain the experimental findings of Shafir and Tversky (1992), according to which subjects cooperate more often in a prisoner's dilemma when they are not told the choice of the other player rather than when they know that the other has cooperated. While players are best reply reasoners when their beliefs about the action of the other is fixed (in line with the ratifiability assumption), they are not in presence of uncertainty.[xiii] Masel (2007) formalises the idea of ADB in a 'Bayesian model of quasi-magical thinking', and shows that in public good games a positive correlation between the players' contributions and their beliefs about the strategy of others can explain cooperative behaviours. Hammond (2009) also defended ADB as a case of 'rational folly', since although players know that their actions cannot directly influence the actions of others, they could be better off if they actually hold that belief: it would therefore be rational for them to hold irrational and false beliefs (see Lecouteux, 2015, propositions 11 and 12, for an evolutionary justification of 'rationally irrational' behaviours).

Furthermore, ADB are not incompatible *per se* with Bayesian rationality. The choices of the players are independent (choosing the strategy that maximises my expected payoff cannot directly influence the choice of your strategy), though the players can believe that the strategy of the other players depend on their own action. They can furthermore *rationalise* this belief – at least when they form their beliefs by simulating the reasoning of other Bayes rational players.[xiv]

Proposition 5. *If there exists a strategy profile $\bar{x} \in X$ that Pareto-dominates a Nash equilibrium $x^* \in X$, then $\bar{x}$ is a SBE.*

Proposition 5 can be seen as a generalisation of Binmore's 'fallacy of the twins', since it means that the players can always rationalise the choice of a profile that Pareto dominates the Nash equilibrium (just as in a Prisoner's Dilemma). According to Binmore (1992, pp. 311-312):

> it is false that rational players can restrict their attention in the Prisoners' Dilemma to the main diagonal of the payoff table […] This would only make sense if the two players did not reason independently. If player I could count on player II reasoning precisely as he reasons, then it would be as though he could force her to choose whichever strategy he found expedient simply by choosing it himself. [But if two rational players] reason in the same way in identical circumstances, it is not because they have no alternative but to think identically: it is because the rational thing to think is the same in both cases.

While we agree that two players who 'reason in the same way in identical circumstances' could play a Nash equilibrium (this is precisely our proposition 4), it is only because the beliefs in their massaged belief hierarchy are action-independent, and not because their reasoning processes are independent. Binmore's definition of rationality in games is indeed that players are best-reply reasoners, while – in line with Kadane and Larkey's initial claim – a Bayesian theory of choice in games should not restrict the definition of admissible strategies to ratifiable choices.

Proposition 5 also suggests that the set of SBE is potentially quite large, and therefore that players who do not converge on the same belief hierarchy during the massaging process are likely to miscoordinate. A Prisoner's Dilemma has for instance two SBE:[xv] mutual cooperation (with the underlying belief that both players cooperate if and only if the other cooperates) and mutual defection (with the belief that both players unconditionally play their Nash strategy). P1 could for instance believe that reciprocity is the norm when playing a Prisoner's Dilemma, and that players are more likely to converge during the massaging process on the beliefs that both players cooperate if the other cooperates. However, if P2 believes that the norm is to coordinate on the Nash equilibrium or to play best reply strategies, then the massaging process will lead her to another belief hierarchy, and we could end up in an asymmetric profile (which cannot be a SBE, since the players did not accurately predict the choice of the other players).

5. Conclusion

We argued in this paper that offering a Bayesian theory of choices in games should integrate a psychological theory of belief formation. Rather interestingly, Harsanyi (1982a, p. 122, emphasis in original) initially recognised that a theory of choice in which the rationality of the players is not certain would require:

> an empirically supported *psychological* theory making at least probabilistic predictions about the strategies people are likely to use […] given the nature of the game and given their own psychological make-up

He also acknowledged that looking for a psychological or normative theory of games are 'very different intellectual enterprises, using very different methodologies as a matter of logical necessity' (Harsanyi, 1982a, p. 122) – this is why Kadane and Larkey (1982b, p. 124) declared 'our differences with Professor Harsanyi are not as profound as might appear'. Harsanyi (1982b, p. 125, emphasis in original) however argued that normative game theory could provide a solid foundation for a descriptive theory of games, since the actual choice of the players is '*either* […] the *correct* move prescribed by normative arithmetic, *or* […] a psychologically understandable deviation from it'. Aumann's reply (in Aumann, 1987) to Kadane and Larkey turned out to be remarkably similar, since the prior beliefs he attributes to the players are an equilibrium of the game. We however argued that Aumann's argument that the essence of game theory is to impose CBR on the set of prior beliefs is problematic, because the identification of the choices compatible with CBR requires solving the game before actually playing it. If players are Bayes rational, the only options available to them are indeed those that maximise their expected payoff; but identifying the set of rational strategies requires defining *ex ante* the beliefs of the players. If Bayesian rationality is common belief, the prior beliefs must be the equilibrium of the game (players cannot indeed simultaneously believe that the others are rational, and that they could play a non-rational strategy). A solution to escape this infinite regress is to investigate the process of belief formation, i.e. how players actually identify the beliefs they attribute to the others.

We argued that game theory did not provide the adequate tools to explain the formation of these subjective beliefs, and suggested introducing the players' capacity of mindreading in game theory. We assumed that players form their beliefs by simulating the reasoning of the other players, and showed that the belief hierarchy of Bayes rational players, when their rationality is common belief (by proposition 1) does not necessarily rule out non-ratifiable choices, and therefore action-dependent beliefs. In presence of uncertainty, a player who simulates the reasoning of the other can indeed take her own choice as an evidence of the choice of the other, because she tends to assume that the reasoning processes of the other are similar to her own reasoning processes.[xvi] We were finally able to derive a solution concept – the subjective belief equilibrium – capturing the mutual Bayes rationality of the players.

An apparent difficulty of our approach is the multiplicity of subjective belief equilibria. An acceptable solution concept should indeed select a restricted number of strategy profiles. However, proposition 5 implies that the set of SBE is quite large, because of the multiplicity of subjective beliefs to which players can converge during the massaging process. A descriptive theory of choice in games should therefore go beyond the mere mathematical representation of games, and not try to give a rational *determinate* solution to games. Pure deductive reasoning is insufficient to give determinate solutions in the majority of games (see e.g. Schelling, 1980[1960], p. 163; and Sugden, 1991, in the context of bargaining games).

Rather than discarding the social, cultural, psychological, and historical background of the players, and look for solutions that rational players could not fail to find, a descriptive theory of games should investigate how those backgrounds affect the formation of individual beliefs (through the definition of focal points and social norms for instance, see e.g. Schelling, 1980[1960]). A descriptive theory of strategic choice should thus rest on works in cognitive and social psychology, so as to characterise the formation of the initial gut feelings of the players, which are determined by personal and social experience (see Scazzieri, 2008, 2011; Bacharach, 1990). Neglecting those factors 'deshumanizes the decision-maker in the opposite direction to the traditional idealization of her powers: instead of exaggerating her resources, it understates them' (Bacharach and Hurley, 1991, p. 3). Colman and Bacharach (1997, pp. 8-9) for instance assume the transparency of *deliberation* and not merely the transparency of *reason*, which is captured through CBR. Various works in social psychology 'have revealed a remarkable degree of consensus in people's understanding of their social environment' (p. 9), and research in *attribution theory* – a theory in social psychology explaining how people attribute second-order beliefs, which is tightly related to the Theory of Mind (see Bacharach, 1986) – has 'shown that the same basic cognitive processes underlie people's predictions and explanations of their own behaviour and that of others' (p. 9). Beyond understanding how the others reason (thanks to the transparency of reason), the individuals can understand how the others discriminate between different ‘rational’ alternatives (thanks to the transparency of deliberation), and therefore can better predict their *actual* choices. In addition to providing a non-tautological explanation of prior beliefs in games, the introduction of the players’ capacity of mindreading, thanks to the Simulation Theory, could therefore constitute the basis of a more general theory of social interactions, based on a genuinely intersubjective theory of behaviour in strategic environments.

Appendix

Proof of proposition 1

Since *i* is uncertain about *j*’s belief, and that *i* attributes her own beliefs to others, we have $B_i(E) \cap U_{ij}(E) \Rightarrow B_{ij}(E)$. Since *i* believes *E* and believes that *j* also believes *E*, *i* believes that *E* is mutual belief: $B_i(E) \cap B_{ij}(E) = B_i\big(MB(E)\big)$. *i* however does not know whether *j* believes that *E* is mutual belief or not. By $SIMB_i$ we have therefore:

$$B_i\big(MB(E)\big) \cap U_{ij}\big(MB(E)\big) \Longrightarrow B_{ij}\big(MB(E)\big) \tag{12}$$

Since *i* believes that *E* is mutual belief, and that *j* also believes it, *i* believes that it is mutual belief that *E* is mutual belief: $B_i\big(MB(E)\big) \cap B_{ij}\big(MB(E)\big) = B_i\big(MB^2(E)\big)$. By continuing the iteration, we find that *i* believes that *E* is common belief.

Proof of proposition 2.

We start by proving $s_{i,j} = s_{i,[k],j}$ for any sequence $[k]$ of players. We consider three players *i*, *j*, and *k* and their respective choice problems:

P_j: *j* must choose a strategy $x_j \in X_j$ so as to maximise her expected payoff

P_k: *k* must form a first order belief $s_{k,j} \in X_j$

P_i: *i* must form a second order belief $s_{i,k,j} \in X_j$

By PM, *j*'s choice function is $f_j(s_{j,-j}) = \underset{x_j \in X_j}{\text{argmax}}\, E\Pi_j(x_j; s_{j,-j}(x_j))$, with:

$$E\Pi_j\left(x_j; s_{j,-j}(x_j)\right) = \sum_{x_i \in X_i} s_{j,-j}(x_{-j}|x_j)\, \Pi_j(x_{-j}, x_j). \tag{13}$$

k's choice function is given by $SIMR_k$: since PM is true for *k*, *k* would also choose the strategy that maximises *j*'s payoff, if she had to choose at her place. We have therefore:

$$\mathcal{C}_k(P_j) = f_k(s_{k,j,-j}), \tag{14}$$

$$\mathcal{C}_k(P_j) = \underset{x_j \in X_j}{\text{argmax}}\, E\Pi_j\left(x_j; s_{k,j,-j}(x_j)\right). \tag{15}$$

Consider now the case of player *i*, who must form a belief about *k*'s belief about *j*'s strategy. If *i* had to form a belief about *j*'s strategy, she would simulate her reasoning (just as *k* in the previous case):

$$\mathcal{C}_i(P_j) = \underset{x_j \in X_j}{\text{argmax}}\, E\Pi_j\left(x_j; s_{i,j,-j}(x_j)\right). \tag{16}$$

i therefore assumes that *k* also simulates *j*'s reasoning (she indeed assumes that *k* has the same reasoning process than hers, i.e. attributing her own reasoning process (PM) to simulate the reasoning of another player). We have therefore:

$$\mathcal{C}_i(P_k) = \underset{x_j \in X_j}{\text{argmax}}\, E\Pi_j\left(x_j; s_{i,k,j,-j}(x_j)\right). \tag{17}$$

Here we can clearly see that $s_{i,j} = s_{i,k,j}$ (condition (i) of proposition 2) if $s_{i,j,-j} = s_{i,k,j,-j}$. In the absence of a prior belief about the beliefs of player *k*, *i* simply attributes her own beliefs to *k* by $SIMB_i$. We thus have $s_{i,j,-j} = s_{i,k,j,-j}$ (*i* assumes that, if she believes $s_{i,j,-j} = x'_{-j}$, then *k* also believes it, i.e. $s_{i,k,j,-j} = x'_{-j}$), and therefore $s_{i,j} = s_{i,k,j}$. We could reproduce the exact same reasoning for higher-order beliefs: player *i* will indeed simulate the reasoning of the succession of players k_1, k_2, etc. which leads her *in fine* to choose the strategy that maximises the expected payoff of *j* ($SIMB_i$ then allows player *i* to attribute her belief about k_1, and then k_2, etc.). This proves part (i) of proposition 2. The proof of part (ii) is similar to the proof of part (i): i indeed considers the beliefs of other players *k* about her own choice, and simulating the maximisation of $E\Pi_i$ in player k's mind, or simulating player *j* simulating the maximisation of $E\Pi_i$ in player k's mind leads to the same outcome if they all share the same higher-order beliefs, which is ensured by $SIMB_i$.

We now turn to the proof of part (iii). For notational convenience, we give the proof for two players: the generalisation to *n* players is conceptually similar but less tractable (we must indeed determine the beliefs of P1 about P2's belief about P3's belief … about player *i*'s strategy, which makes the notations quite heavy although the resolution is based on a fixed point argument which is conceptually similar for *2* and *n>2* players). PM means that player *i* chooses the strategy that maximises her expected payoff given her beliefs about *j*'s strategy (that may depend on *i*'s strategy). By simulation, *i* assumes that PM is true for *j* (it indeed corresponds to her choice function): *i* therefore believes that *j* chooses her strategy so as to maximise her expected payoff, given *j*'s beliefs about *i*'s strategy (that may also depend on j's strategy). If player *i* plays a strategy $\hat{x}_i$, she believes that *j* plays:

$$s_{i,j}(\hat{x}_i) = \underset{x_j \in X_j}{argmax} \left[\sum_{x_i \in X_i} s_{i,j,i}(x_i | x_j)\, \Pi_j(x_i, x_j) \right] \tag{18}$$

with $s_{i,j,i}: X_j \longmapsto \Delta(X_i)$ *i*'s belief about *j*'s belief about *i*'s strategy. *i* must therefore anticipate how *j* forms her belief about her actions. Since she attributes her own reasoning process to *j*, she assumes that *j* forms her belief by simulating *i*'s reasoning process. Since *i* may believe that her choice could be correlated with *j*'s choice, she believes that *j* also believes that their choices could be correlated (by SIMB$_i$). If player *j* plays a strategy $\hat{x}_j$, then *i* believes that *j* believes that *i* plays:

$$s_{i,j,i}(\hat{x}_j) = \underset{x_i \in X_i}{argmax} \left[\sum_{x_j \in X_j} s_{i,j,i,j}(x_j | x_i)\, \Pi_i(x_i, x_j) \right] \tag{19}$$

with $s_{i,j,i,j}: X_i \mapsto \Delta(X_j)$ *i*'s belief about *j*'s belief about *i*'s belief about *j*'s strategy. By simulation, this function is also defined as the best reply to a higher order belief, and so on.

Note that we have only investigated *i*'s belief about *j*'s action for a given strategy $\hat{x}_i$: the same operation should then be done for each possible strategy in X_i (and at each step, for all the strategies $\hat{x}_j \in X_j$ of player j). A more concise notation of the problem would be the following:

$$s_{i,j}(x_i) = \underset{f_j \in F_j}{argmax} \left[\sum_{x_i \in X_i} s_{i,j,i}\left(f_j(x_i)\right) \Pi_j(x_i, x_j) \right], \tag{20}$$

$$s_{i,j,i}(x_j) = \underset{f_i \in F_i}{argmax} \left[\sum_{x_j \in X_j} s_{i,j,i,j}\left(f_i(x_j)\right) \Pi_i(x_i, x_j) \right] \tag{21}$$

with $F_j = \{f_j: X_i \mapsto \Delta(X_j)\}$ the set of functions associating a probability distribution $p_j \in \Delta(X_j)$ for a strategy $\hat{x}_i \in X_i$. We however know that $s_{i,j,i,j} = s_{i,j}$ at the MBH: this means that $s_{i,j}$ and $s_{i,j,i}$ must be a mutual best reply. When massaging her beliefs, *i* is therefore looking for a vector of conditional probability distributions $\{f_j^*, f_i^*\}$ that simultaneously maximises the expected payoff of both players, i.e. such that:

$$\sum_{x \in X} s^*(x)\Pi_i(x) \geq \sum_{x \in X} s'^{(x)\Pi_i(x)}, \qquad \forall s' \in \Omega\left(f_j', f_i^*\right), \forall f_j' \in \Phi_i, \forall i \in N, \tag{22}$$

With $s^* \in \Omega\left(f_j^*, f_i^*\right)$. The probability distribution s^* representing the vector of conditional probability distributions $\left\{f_j^*, f_i^*\right\}$ verifies condition (iii).

Proof of proposition 3.

Consider the unconditional beliefs $s_{i,j} = p_j^*$ and $s_{i,j,i} = p_i^*$ $\forall i, j \in N$ (and suppose that the higher-order beliefs of *i* correspond to her 1st and 2nd order beliefs). Player *i* believes that all the other players unconditionally play their Nash (mixed) strategy, and believes that all the players *j* believe that she unconditionally plays her Nash strategy. Even if player *j* adopts another conditional distribution, she knows that the strategy of the others will remain the same: the highest payoff player *j* can get is therefore the one induced by her best reply to p_{-j}^*, i.e. her Nash strategy p_j^*. The belief hierarchy S_i^* such that all the players unconditionally play their Nash strategy is therefore a massaged belief hierarchy.

Proof of proposition 4.

The proposition is a corollary of proposition 3: if $s_{i,j}^* = x_j^*$ and $s_{i,j,i}^* = x_i^*$, $\forall i, j \in N$, with $x^* \in X$ the Nash equilibrium of G, then the belief hierarchy generated by s^* is a MBH. The optimal strategy for each player (the strategy that maximises their expected payoff) is their Nash strategy by construction, and – since the equilibrium is in pure strategies – they accurately predict the choice of the other players.

Proof of proposition 5.

Let $x^* \in X$ denote a Nash equilibrium of G and $\bar{x} \in X$ a strategy profile that is Pareto superior to p^*, i.e. a profile such that:

$\Pi_i(\bar{x}) \geq \Pi_i(x^*), \ \forall i \in N,$

with at least one strict inequality. Consider the following beliefs:

$$\begin{cases} s^*_{i,j}(x_i \neq \bar{x}_i) = x^*_j, \ \ s^*_{i,j,i}(x_j \neq \bar{x}_j) = x^*_i, & \forall i,j \in N, i \neq j \quad (23) \\ s^*_{i,j}(x_i = \bar{x}_i) = \bar{x}_j, \ \ s^*_{i,j,i}(x_j = \bar{x}_j) = \bar{x}_i, & \forall i,j \in N, i \neq j \quad (24) \end{cases}$$

Condition (23) means that, when the players choose a strategy different from $\bar{x}_i$, they believe that all the other players unconditionally play their Nash strategy. However, condition (24) means that, when they play $\bar{x}_i$, they all believe that all the others play their part of the profile $\bar{x}$. Player *i* should therefore play $\bar{x}_i$ to maximise her expected payoff (condition (i)): if *i* chooses another strategy, she indeed believes that all the other plays their Nash strategy, and the highest payoff for *i* in this case is her Nash payoff. Since they all reach a higher payoff at $\bar{x}$, they should all play their part of the profile. Furthermore, by construction of their beliefs, they predict well the choice of the other players (condition (ii)). Lastly, the belief hierarchy generated by s^* is a MBH: changing one's conditional probability can either lead *i* to play $\bar{x}_i$ (in which case the payoff is the same) or a different strategy $x_i \neq \bar{x}_i$, which induces all the other players to play their Nash strategy (in which case player *i*'s payoff is bounded by her Nash payoff, which is lower than the initial payoff). $\bar{x}$ is well a subjective belief equilibrium.

References

Anscombe, F., & Aumann, R. (1963). A definition of subjective probability. *Annals of Mathematical Statistics, 34*(1), 199–205.

Arnold, B. C., & Press, J. S. (1989). Bayesian estimation and prediction for Pareto data. *Journal of the American Statistical Association*, *84*(408), 1079-1084.

Aumann, R. J. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica, 55*(1), 1–18.

Aumann, R., & Brandenburger, A. (1995). Epistemic conditions for Nash equilibrium. *Econometrica*, *63*(5), 1161-1180.

Aumann, R. J., & Dreze, J. H. (2008). Rational expectations in games. *The American Economic Review*, *98*(1), 72–86.

Bacharach, M. (1986). The problem of agents' beliefs in economic theory. In M. Baranzini & R. Scazzieri (Eds.), *Foundations of Economics*. (pp. 175-203). Oxford, UK: Blackwell.

Bacharach, M. (1990). Commodities, language, and desire. *The Journal of Philosophy*, *87*(7), 346-358.

Bacharach, M., & Hurley, S. (1991). Introduction. In M. Bacharach & S. Hurley (Eds.), *Issues and advances in the foundations of decision theory* (pp. 1-38). Oxford, UK: Blackwell.

Battalio R., Samuelson L., & Van Huyck J. (2001). Optimization incentives and coordination Failure in Laboratory Stag Hunt Games. *Econometrica*, *69*(3), 749-764.

Bernheim, D. (1984). Rationalizable strategic behaviour. *Econometrica, 52*(4), 1007–1028.

Binmore, K. G. (1992*). Fun and games: A text on game theory*. Lexington, MA: Heath.

Binmore, K. G. (2006). Rational decisions in large worlds. Lecture given at 2006 ADRES Conference (Marseille, France). http://discovery.ucl.ac.uk/14433/1/14433.pdf

Binmore, K. G. (2009). *Rational decisions*. Princeton, NJ: Princeton University Press.

Board, O. (2006). The equivalence of Bayes and causal rationality in games. *Theory and Decision*, *61*(1), 1-19.

Böge, W., & Eisele, T. (1979). On solutions of bayesian games. *International Journal of Game Theory*, *8*(4), 193-215.

Bonanno, G., & Nehring, K. (1999). How to make sense of the common prior assumption under incomplete information. *International Journal of Game Theory, 28*(3), 409-434.

Brandenburger, A. (2010). Origins of Epistemic Game Theory, in *Epistemic Logic: Five Questions,* ed. by Vincent Hendricks and Olivier Roy, *Automatic Press,* 59-69.

Colman, A. M., & Bacharach, M. (1997). Payoff dominance and the Stackelberg heuristic. *Theory and Decision, 43*(1), 1-19.

Foster, D., & Young, P. (1990). Stochastic evolutionary game dynamics∗. *Theoretical population biology, 38*(2), 219-232.

Friedell, M. F. (1969). On the structure of shared awareness. *Behavioral Science, 14*(1), 28–39.

Gardner, M. & Nozick, R. (1974). Reflections on Newcomb's problem: a prediction and free-will dilemma. *Scientific American*, *230*(3), 102-109.

Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford, NY: Oxford University Press.

Goldman, A., & Shanton, K. (2012). The case for simulation theory. In A. Leslie & T. German (Eds.), *Handbook of 'Theory of Mind'* (pp. 202-204), chapter. Mahwah, NJ: Erlbaum.

Gul, F. (1998). A comment on Aumann's Bayesian view. *Econometrica*, *66*(4), 923-927.

Hammond, P. J. (2009). Isolation, assurance and rules: can rational folly supplant foolish rationality?. In K. Basu & R. Kanbur (Eds.), *Arguments for a better world: Essays in honor of Amartya Sen. Volume I: Ethics, welfare, and measurement.* Oxford, NY: Oxford University Press.

Harsanyi, J. C. (1967-1968). Games with incomplete information played by 'Bayesian' players, I-III, *Management Science, 14*(3-5-7), pp. 159-182, 320-334, 486-502.

Harsanyi, J. C. (1982a). Comment – subjective probability and the theory of games: Comments on Kadane and Larkey's paper. *Management Science*, *28*(2), 120-124.

Harsanyi, J.C. (1982b). Rejoinder to professors Kadane and Larkey. *Management Science, 28*(2), 124-125.

Haruvy, E., Stahl, D. E., & Wilson, P. W. (1999). Evidence for optimistic and pessimistic behavior in normal-form games. *Economics Letters*, *63*(3), 255–259

Hausman, D. M. (2000). Revealed preference, belief, and game theory. *Economics and Philosophy*, *16*(1), 99–115.

Hausman, D. M. (2012). Preference, choice, value and welfare. NY: Cambridge University Press.

Hédoin, C. (2016). Community-Based Reasoning in Games: Salience, Rule-Following, and Counterfactuals. *Games*, *7*(4), 36.

Heidl, S. (2016). *Philosophical problems of behavioural economics*. Abingdon: Routledge.

Jeffrey, R. C. (1990). *The logic of decision*. Chicago: University of Chicago Press.

Kadane, J. B., & Larkey, P. D. (1982a). Subjective probability and the theory of games. *Management Science, 28*(2), 113–120.

Kadane, J. B., & Larkey, P. D. (1982b). Reply to Professor Harsanyi. *Management Science*, *28*(2), 124.

Kirman, A., & Teschl, M. (2010). Selfish or selfless? The role of empathy in economics. *Philosophical Transactions of Royal Society B,* 365: 303–317.

Kramer, R. M., & Brewer, M. B. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology*, *46*(5), 1044-1057.

Lecouteux, G. (2015). Reconciling Normative and Behavioural Economics (PhD thesis), École Polytechnique, Palaiseau, France.

Lehtinen, A. (2011). The revealed-preference interpretation of payoffs in game theory. *Homo Oeconomicus, 28*(3), 265–296

Levi, I. (1998). Prediction, Bayesian deliberation and correlated equilibrium. In W. Leinfellner & E. Köhler (Eds.), *Game Theory, experience, rationality: foundations of social sciences, economics and ethics in honor of John C. Harsanyi* (pp. 173–185). Dordrecht: Springer.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York, NY: Wiley.

Mariotti, M. (1996). The decision-theoretic foundations of game theory. In K. Arrow, E. Colombatto, M. Perlman, & C. Schmidt (Eds.), *The Rational Foundations of Economic Behaviour* (pp. 133-148). Basingstoke: Macmillan Press Ltd.

Masel, J. (2007). A Bayesian model of quasi-magical thinking can explain observed cooperation in the public good game. *Journal of Economic Behavior & Organization*, *64*(2), 216-231

Morris, S. (1995). The common prior assumption in economic theory. *Economics and Philosophy, 11*(02), 227–253.

Perea, A. (2014). From classical to epistemic game theory. *International Game Theory Review*, *16*(1), 1440001-1–1440001-22.

Sanghvi, A.P., & Sobel, M. J. (1976). Bayesian games as stochastic processes. *International Journal of Game Theory*, *5*(1), 1-22.

Savage, L.J. (1954). *The Foundations of Statistics*. New York, NY: Wiley.

Scazzieri, R. (2008). Coordination, context and patterns of Reasoning. In M. C. Galavotti, R. Scazzieri and P. Suppes (Eds.), *Reasoning, Rationality and Probability* (pp. 187-207). Chicago: CSLI Publication, Chicago University Press.

Scazzieri, R. (2011). Similarity and uncertainty. In M. Dall'aste Brandolini Silva and R. Scazzieri (Eds), *Fundamental Uncertainty. Rationality and Plausible Reasoning* (pp. 73-103). Basingstoke: Palgrave Macmillan.

Schelling, T. C. (1980[1960]). The Strategy of Conflict. Cambridge, MA: Harvard University Press.

Segal, N.L., & Hershberger, S.L. (1999). Cooperation and competition between twins: findings from a Prisoner’s Dilemma game. *Evolution and Human Behavior, 20*(1), 29-51.

Shafir, E., & Tversky, A. (1992). Thinking through uncertainty: non consequential reasoning and choice. *Cognitive Psychology*, *24*(4), 449-474.

Shanton, K., & Goldman, A. (2010). Simulation theory. *WIREs Cognitive Sciences*, *1*(4), 527-538. doi: 10.1002/wcs.33. Retrieved from: http://onlinelibrary.wiley.com/doi/10.1002/wcs.33/full

Singer, T., & Fehr, E. (2005). The neuroeconomics of mind reading and empathy. *American Economic Review*, *95*(2), 340 – 345.

Sugden, R. (1991). Rational bargaining. In S. Hurley & M. Bacharach (Eds.), *Foundations of Decision Theory* (pp. 294-316). Oxford, UK: Basil Blackwell.

Sugden, R. (2001). The evolutionary turn in game theory. *Journal of Economic Methodology*, *8*(1), 113-130.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Belmont, CA: Brooks/Cole.

Vanberg, V. J. (2008). On the economics of moral preferences. *American journal of Economics and Sociology*, *67*(4), 605-628.

Von Neumann, J. & Morgenstern, O. (1944). *The theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Von Neumann, J. & Morgenstern, O. (1947). *The Theory of games and economic behavior*. 2nd rev. ed. Princeton, NJ: Princeton University Press.

---

[i] Luce and Raiffa (1957, p. 306) already suggested that players should choose the best act given a prior probability distribution over the strategies of the other players, generated by the strategic properties of the game and the psychological information about the others. Kadane and Larkey (1982a, p. 115) also mention related works by Böge and Eisele (1979) and Sanghvi and Sobel (1976).

ii The initial paper by Kadane and Larkey was commented by Harsanyi in the same issue of Management Science, followed by a reply by Kadane and Larkey and a rejoinder by Harsanyi.

iii We assume here for simplicity that the preference ordering corresponding to the mentalistic interpretation (one's material payoff) is the same than the ranking of monetary gains: a more precise definition would however include other-regarding preferences (in the sense of Vanberg, 2008), as preferences over outcomes rather than actions), such as sympathetic concerns for others, rather than restricting mentalistic preferences to the player's self-interest.

iv Other issues are discussed in the literature such as (i) the subjective identification of outcomes: explaining the choice of the individuals – and identifying whether they are consistent or not – requires knowing how the agent identifies the different outcomes (see Heidl, 2016, pp. 30-32), and (ii) the identity of choice and preference: although preferences in games are defined over outcomes and strategy profiles, individual choice actually consists in a single element of a strategy profile (Lehtinen, 2011, p. 275).

v Goldman (2006) considers that the role of pretence, which is the core of the simulation routine, is not restricted to mental states. It intervenes either for processes (i.e. a decision-making mechanism) or for its inputs (i.e. beliefs, desires, etc.) and outputs (i.e. the decision). In this paper, we will only consider that players simulate the process of decision-making and its inputs (when players have neither information nor specific prior belief about the beliefs of others).

vi Note that the belief about P1's action that P1 attributes to P2 is not justified for now: we will determine in section 3 what beliefs P1 could attribute to P2 if she also believes that P2 is a Bayes rational player.

vii By 'consistency', we mean the internal consistency between the player's beliefs about the rationality of the others and her belief about their actions. The beliefs that you will play a strictly dominated strategy and that you are a best reply reasoner are for instance inconsistent, and requires me revising my beliefs (the 'massaging process' corresponds to this phase of self-introspection during which I revise my beliefs until they become consistent).

viii Battalio, Samuelson and Van Huyck (2001) suggests for instance that the probability of choosing the risk-dominant equilibrium tends to increase with the optimisation premium (the difference between the payoff of the best response and the inferior response – which is for instance larger when $\Pi_1(A, A)$, $\Pi_1(A, B)$, and $\Pi_1(B, B)$ are of a similar magnitude, while being significantly higher than $\Pi_1(B, A)$). However, the subjects in their experiment were probably not aware that their choice was actually influenced by this premium (for the simple reason that they probably did not explicitly calculate the premium before choosing).

ix Friedell does not use the expression 'common belief' but 'common opinion'. The two concepts are however mathematically identical (see Perea, 2014, p. 11).

x A possible game-theoretic definition of condition (iii) would be to interpret a MBH as the Nash equilibrium of a game of 'beliefs' in which the individual sets of strategies are not X_i as in the initial game, but a set of conditional strategies $Y_i = \{f: X_{-i} \mapsto X_i\}$, i.e. strategies conditioned on the strategy of the other players (such that 'play C if the other plays C, and D otherwise'). It would not however be possible to properly define this game in normal form, because the set of conditional strategies depends on the set of conditional strategies of the other players: since two conditional probabilities $P\ (X_1|X_2)$ and $P\ (X_2|X_1)$ cannot always define a distribution $P(X)$, some conditional strategies are incompatible, i.e. it would not be possible to define a probability distribution representing the choice of the players (meaning that the payoff of the players would not be defined for some strategy profiles).

xi Bernheim (1984, p. 1014) explicitly states a similar (but stronger) condition of uncorrelated beliefs: 'the choices of any two agents are by definition independent events [...] Consequently, I restrict players to have uncorrelated probabilistic assessments of their opponents' choices'. However, as pointed by Levi (1998, footnote 9) and discussed in footnote xii, Aumann (1987) claims that Bayes rational choices should be ratifiable: this claim then remained implicit in the subsequent literature in epistemic game theory.

xii Note that Aumann (1987) defends the idea that beliefs can be correlated (suggesting that the choices of the player could be dependent), but still defines Bayes rational choices as ratifiable choices. His justification is that the choice of the individual is a two-step process (Aumann 1987, pp. 3-4): given an initial probability distribution over the set of profiles, the individual is firstly informed of her strategy (which gives her a belief about the strategy of the others – this belief can therefore be conditional on the strategy suggested to her), and then independently chooses her best reply given her beliefs. Player i can choose a new strategy without affecting her belief, because the actions of the other players are 'fixed' to the state of the world revealed to the player in the first stage. This interpretation however requires the existence of a public signal – e.g. a social norm – to inform all the players of the profile selected in the first stage: actions are therefore independent, while beliefs may be correlated by the public signal. This two-step structure was initially suggested by Harsanyi (1967-1968), but his argument was that 'nature' selects in a first stage the type of the players (their utility functions) only, and not their strategies.

xiii Similar predictions are found with Newcomb's problem by Gardner and Nozick (1974) and Shafir and Tversky (1992): around 2/3 of the individuals chose to take only one box, and not to play the dominant strategy of taking the two boxes (as if they believed that their action could actually affect the probability of getting a high outcome in the 'risky' box).

[xiv] Board (2006) for instance show that Aumann's (1987) notion of Bayes rationality is equivalent to causal rationality (corresponding to our equation (11)) only if we add a condition of causal independence. Our point is that ST could question the players' belief in the causal independence of their actions. Indeed, if causal independence were common belief among the players, then our proposition 5 would not hold (see Hédoin, 2016, pp. 11-15, for a related point)

xv We can indeed show that all the players get at least their maximin payoff at a SBE, which excludes the asymmetric profiles from the set of possible SBE. If player i has less than her maximin, then the underlying belief hierarchy cannot be a MBH: it would indeed be in the interest of i to unconditionally play her maximin strategy (the other players cannot therefore rationally believe that i would play her part of the SBE).

xvi We can mention Segal and Hershberger's (1999) experiment on cooperation between twins to support our hypothesis that 'similar' individuals are more likely to form ADB (they indeed find that monozygotic (identical) twins are more likely to cooperate in a prisoner's dilemma compared to dizygotic twins). Furthermore, several studies on social identity theory (Tajfel and Turner, 1979) suggest that players who think of themselves as members of a common group are more likely to cooperate in social dilemmas (e.g. Kramer and Brewer, 1984). A possible explanation of those results is that 'similar' individuals (whether it be socially or genetically) are more likely to believe that their actions are correlated, and therefore to form ADB (which could then explain the higher rates of cooperation in the prisoner's dilemma).

DOCUMENTS DE TRAVAIL GREDEG PARUS EN 2017
GREDEG Working Papers Released in 2017

2017-01 Lauren Larrouy & Guilhem Lecouteux
Mindreading and Endogenous Beliefs in Games