

CHALLENGING STANDARD NON-COOPERATIVE GAME THEORY? FROM BACHARACH'S "VARIABLE FRAME THEORY" TO "TEAM REASONING"

Documents de travail GREDEG
GREDEG Working Papers Series

LAUREN LARROUY

GREDEG WP No. 2014-17

<http://www.gredeg.cnrs.fr/working-papers.html>

Les opinions exprimées dans la série des **Documents de travail GREDEG** sont celles des auteurs et ne reflètent pas nécessairement celles de l'institution. Les documents n'ont pas été soumis à un rapport formel et sont donc inclus dans cette série pour obtenir des commentaires et encourager la discussion. Les droits sur les documents appartiennent aux auteurs.

*The views expressed in the **GREDEG Working Paper Series** are those of the author(s) and do not necessarily reflect those of the institution. The Working Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate. Copyright belongs to the author(s).*

Challenging Standard Non-Cooperative Game Theory? From Bacharach’s “Variable Frame Theory” to “Team Reasoning”

Lauren Larrouy*

GREDEG Working Paper No. 2014-17

(very preliminary)

Abstract

The paper purports to stress how the two major contributions of Bacharach: “Variable Frame Theory” (VFT) and “Team Reasoning” (TR) improve *Standard Non-Cooperative Game Theory* in some relevant aspects which I point out. The aims are to show: (i) how Bacharach respectively justifies coordination and cooperation within these theories, and (ii) how these improvements in both VFT and TR involve a new conception of players and their rationality. I underline how coordination and cooperation rely on contextual and social determinants, which challenge and even contradict some pillars of standard individual rationality in terms of subjective expected utility, in games. Even if Bacharach’s conceptual and methodological choices within these theories induce numerous difficulties, I try to show that Bacharach’s work underlines some of the implications and related problems induced by the mere foundations of standard non-cooperation game theory.

Key Words: non-cooperative game theory, framing, salience, focal points, Team Reasoning, coordination, cooperation, individual rationality, social rationality, collective rationality.

JEL Codes: B21, B41, C72, D03, D79

1. Introduction

In its most basic characterization, standard non-cooperative game theory formalizes *contexts of strategic interactions* without communication. Basically, it means that the issue of interactions relies on the combination of the interacting agents’ decisions and then actions. Accordingly, players have to take into account other players’ possible actions. They have to form beliefs about others’ decisions and beliefs. This *specific kind of uncertainty* entails that game theory requires rigorous principles defining (i) players’ decisions (Bacharach, 1976: vii)¹ and (ii) the epistemic requirements of players’ decisions (Colman, 2003).

* GREDEG CNRS, University of Nice Sophia Antipolis, lauren.larrouy@gredeg.cnrs.fr

¹ “[A] game is an activity of two or more human beings whose outcomes matters to each of them, and depends in a *prescribed way* on the actions of all of them” (Bacharach, 1976: 2; my emphasis)

Game theorists still commonly prescribe to von Neumann and Morgenstern's conception of individual rationality (Bacharach, Hurley, 1993: 3-4; Giocoli, 2003), i.e. "an objective definition of rational behavior that could guide a player's choices in a game independently of his/her psychology and opinion on the others players' psychology" (Giocoli, 2003: 13). They "derive a theory of rational play in games from one of rational individual decision-making [...] which we may call *individualism* in game theory" (Bacharach, Hurley, 1993: 3-4). That is why players' rationality in games presupposed – and still presupposes – an internal consistency of choices and maximization of their subjective expected utility². von Neumann and Morgenstern's conception of game theory as a "tool-box of powerful analytical methods" (Giocoli, 2003: 13) and of rationality must be considered alongside with an *axiomatic approach*³ which is still reckoned as an untouchable statement of the foundations of standard game theory. Game theory became "an internally closed procedure which operates according to fixed rules known by all mathematicians [and presently by all game theorists]" (Giocoli, 2003:15; quoting von Neumann 1983 [1931]: 61-62). As a result, the premise that mathematics is "the *universal language*" (*ibid.*,16) translates in game theory. Following von Neumann and Morgenstern, game theorists assumed – and again, still assume – that every player can do what they are capable of doing as game theorists, i.e. understanding mathematics to draw valid conclusions from a mathematical structure⁴. It implies that all of the relevant information needed for players to make their decision is contained within the mathematical structure of games. What stems from outside the game does not matter.

As Bacharach (1976:1) emphasizes "[this] methodology is one of abstraction and assimilation". In fact, standard non-cooperative game theorists resort to a representative agent (Davis, 2011); players are homogeneous^{5,6} and their rationality is perfect. Standard game theory remains anchored to the "principle of rational determinacy" (Sugden, 1991), which means that there is only one way to play, one way to decide and to act in all the possible circumstances, i.e., in all types of games. According to Bacharach (1976:1) this conception supposes a 'perfect agent' cognitively over-determined. This methodology encounters numerous difficulties. Even with perfectly rational agents, some problems, materialized *via* the well-known *indeterminacy problem*, remain. Coordination games⁷, in which several equilibria compete, cannot be resolved. It is impossible to derive a solution (i) from the rules of the game, (ii) the agents' knowledge and (iii) the rationality postulate (those things being of common knowledge). However, this type of interactions is recurrent in real life, and individuals are perfectly able to solve it (Schelling, 1980[1960]; Metha et al., 1994). In addition, recurrent deviations from perfect rationality become increasingly obvious with the rise of behavioral

² At the origin consistency of choice and maximization was considered as two distinct grounds of decision theory (see, e.g. Sent [2013]): "Rationality, history of the concept" in the *New Palgrave Dictionary of Economics*). I shall not discuss the debates with respect to the semantic and the foundations of rationality within the decision theory paradigm, as they are beyond the scope of this paper.

³ "[T]he [axiomatic] method consists in the implementation of the following steps: (1) state a list of technical terms or symbols that are chosen as undefined terms: these are the primitives of the theory; (2) lay down a set of unproved statements about the primitives terms: these are the axioms of the theory; (3) verify that the axioms satisfy as a set the requirements of independence, completeness and consistency; and (4) deduce from the axioms all other statements – or theorems in the theory. The procedure is objective, so everyone can reproduce it to check the derivation of a certain result" (Giocoli, 2003:15).

⁴ This statement partly entails a well-known problem of *indeterminacy*, with which both the theorists and the players are consequently faced.

⁵ "The impossibility of accommodating heterogeneous agents is today an acknowledged drawback of standard neoclassical models" (*ibid.*, 20)

⁶ In a way, we can consider that game theory escapes from this pitfall by entailing imperfect information. Nonetheless, players remain fundamentally depicted as representative agents. Ultimately, they draw the same inferences and the same conclusions with common information; they can only reach different conclusions if they have different information. As soon as information is revealed they update their beliefs and can converge to the same conclusions. And they still have the same objective: maximizing.

⁷ Basically, in Coordination games there is no dominant strategy and multiple equilibria which can be Pareto equivalent or not.

economics. Progress within neurology and neuro-economics, as well as psychology, show us that the major part of our decision is not explained by rationalization and deliberation but by psychological aspects, reflexes, and emotions. Individuals are generally inclined to follow norms and social rules, etc. Besides, when we interact with others (in strategic contexts or not) some of their characteristics like their identity eventually matter⁸. Consequently, there are some relevant aspects of players' choices which are not embraced by standard rationality.

For all these reasons, the confessed objective of von Neumann and Morgenstern to set up game theory as an objective and deductive framework, largely applied to social science, unquestionably induces several limits (see, e.g. Colman, 2003). Therefore, with the current developments in *Behavioral Economics*, we are facing an attempt to diminish its level of abstraction (Colman, 2003). I shall try to show that Bacharach has a vision of game theory and especially standard non-cooperative game theory on which he concentrates his work, which is well in line with this new trend. He emphasizes how this dualism between a normative and positive account of game theory, directly inherited from von Neumann and Morgenstern, is harmful for game theory and its relevance in the economic world. He claims that game theory "is going so far, abstracting too much, [and] confusing the essences of quite disparate things" (Bacharach, 1976: 1). Nevertheless, the purpose of this paper is not to dwell on the widespread criticism addressed against standard non-cooperative game theory, quite the opposite it emphasizes how game theory can be at the same time highly powerful and yet limited. Certain enrichments of standard non-cooperative game theory seem hard to reconcile with its fundamentals without endangering the validity of the new theories or models proposed within this framework. It is from this perspective that game theory is a powerful theory.

Bacharach's approach to game theory is first foremost focused on two fundamental problems: (i) the "hyperrationality assumptions" and (ii) the indeterminacy (Bacharach, 1994: 7). What crystallizes all the *possible* advances he proposes – with respect to the problems and failures which game theory faces – can be summarized by two dimensions:

- i. The inclusion of players' frames within games in his "Variable Frame Theory" (VFT) – where players' subjective representations of the game become different from the game theorists' objective presentations (Bacharach, 1991, 1993, 1997, 2001a; Bacharach, in Gold and Sugden, 2006; Bacharach, Bernasconi, 1997; Bacharach, Stahl, 2000),
- ii. The conceptualization of a collective mode of reasoning in his work on "Team Reasoning" (TR), (Bacharach, 1995, 1997, 1998, 1999, 2001a, 2001b; Bacharach, in Gold, Sugden, 2006).

I shall highlight how they can be considered as the major contributions of Bacharach regarding the debates I mentioned, although his work is generally characterized by an important conceptual and methodological as well as analytical diversity. Both in VFT and TR, Bacharach takes into account that the way players represent their decision problem matters; in other words their *frames* matter. Then, as frames are necessarily incomplete (1991, 1993, 1997, 2001) for Bacharach, because players are not "conceptually omniscient" (1991:29) and because framing is prior to reasoning (2001: 5), "the hyperrationality assumption" vanishes. In line with the argument concerning the avoidance of a perfect individual rationality, Bacharach argues that the conception of players in standard non-cooperative game theory is a misleading path which has shown its limits (again, referring to the well-known indeterminacy problem). According to him what goes wrong in game theory is:

"In games as traditionally specified the attitude data are limited to preferences over the game's outcomes, plus knowledge of the rules of the game and of the other players'

⁸ See, e.g. Colman (2003) or Gintis (2007) who gather diverse and complementary fields of inquiry on these multiple possible deviations and their theoretical implications.

rationality, preferences, and similar knowledge. [...] So we may not use the fact that every real player has the general knowledge her culture gives her, such as knowledge of which arrangements are salient or traditional in that culture and so provide coordination. The limitation to attitude data which omit such knowledge dehumanizes the decision-maker in the opposite direction to the traditional idealization of her powers: instead of exaggerating her resources, it understates them” (Bacharach, Hurley, 1991: 3).

That is why rationality should be understood as a *valid* mode of reasoning within a *realistic* conception of players. In fact, Gold and Sugden (2006: 9) highlight how “it is fundamental to Bacharach’s methodology that an analysis of valid reasoning about a game can help to explain the behaviour of real players”. Consequently, in order to show how players’ rationality changes, I shall point out how rationality in Bacharach’s work is linked to his notion of what a real player is.

Colman (quoting Binmore) accurately introduces a current statement within game theorists:

“Game theorists of the strict school believe that their prescriptions for rational play in games can be deduced, in principle, from one-person rationality considerations without the need to invent collective rationality criteria – provided that sufficient information is assumed to be common knowledge” (Colman, 2003: 143; Binmore, 1994: 143)

I will emphasize that Bacharach’s VFT and TR, like Sugden’s TR (1993, 2000, 2005, 2007), turn that widely held tendency denounced by Colman upside down. Nevertheless, Bacharach’s specific methodological choices raise numerous difficulties from a theoretical and analytical point of view when compared with Sugden’s account of TR which does not exhibit such difficulties. What we feel in Bacharach’s work is a gnawing pain between objective and subjective features of players to justify both coordination and cooperation in a standard non-cooperative framework. Notwithstanding, Bacharach begs worthwhile questions.

The paper is structured as follows. The first section describes Bacharach’s VFT and its evolution. The purpose is to emphasize the dualism he faces between a normative account of game theory and a descriptive one. I shall try to show how the introduction of subjective characteristics in players’ decisions threatens the possibility of building a game and hence to describe interaction within a game theoretical framework. The second section presents Bacharach’s TR and again how its evolutions highlight some difficulties. If VFT begs the question of the common basis required for individuals to interact within a game theoretic framework, TR begs the question of “what is an individual” when she refers to a group with whom she is interacting and she considers herself with respect to this group. The last section details why Bacharach’s conception of players must be considered in line with two other types of economic rationality, i.e. social and collective rationality. Then I specify upon what basis players’ rationality becomes social and collective. Although we understand what social and collective rationality intuitively mean, analytically it is more obscure. For this reason, I shall point out what are the analytical implications of these two forms of rationality.

2. Bacharach’s “Variable Frame Theory” and coordination

Both methodological and conceptual aspects of Bacharach’s work on “Variable Frame Theory” (VFT) evolve. Nonetheless some grounds allow us to identify one recurrent purpose and a “general” theoretic unity within the different contributions. The purpose of this theory is to integrate Schelling’s ‘focal points’ in a game theoretical framework (Bacharach, 1991: 5; Bacharach, in Binmore et al., 1993: 270; Bacharach, Bernasconi, 1997: 2); which for Bacharach

means that the way players' coordinate on focal points has to be drawn from *rigorous theoretical principles*, i.e, from a *valid*⁹ theory of games.

It is now widely acknowledged that focal points are of a particular interest for standard game theory. They challenge the problem of *equilibrium selection* within games. Indeed, focal points and salience are considered to be one path within the *Refinement Program* to escape from the problem of *indeterminacy* (Sugden, 2001: 116). Accordingly game theorists showed a major interest in this concept (Metha, Starmer and Sugden, 1994; Sugden, 1995; Casajus, 1998; Colman, 1997; Janssen, 2001, 2006; Sugden and Zamarron, 2006). Nonetheless, focal points and salience never became a *game theoretical principle* (Sugden, 2001: 116, Innocenti, 2007; Bacharach, 1991, 1993; Bacharach and Bernasconi, 1997).

Bacharach's methodology to incorporate 'focal points' in a game theoretical framework was to integrate players' frames within games. He defines frames as follows: "[a] frame is a set of *concepts* or *predicates* an agent uses in thinking about the world" (Bacharach, 2001a: 1). More specifically, "a players' frame is, most simply, the set of variables she uses to conceptualize the game" (Bacharach, 1997: 4). Bacharach uses framing in a very specific way, which is quite unusual within the literature on "framing effects"^{10,11}. However, regardless the case, framing puts an emphasis on the fact that 'the act' of choosing requires two steps: (i) the *conceptualisation of the decision problem*, i.e. "the choice problem is 'edited' or 'framed'" (Klaes, 2008: 215); and (ii) the *evaluation of the different options the agent faces* with respect to the way she framed her decision problem: "the framed alternatives are then evaluated" (*ibidem*).

In this section my attempt is mainly to emphasize how and why Bacharach's VFT is different from standard non-cooperative game theory. I shall present the common features of Bacharach's models of VFT in the first subsection in order to emphasize the methodological and conceptual differences between them in the second subsection. They show Bacharach's difficulty in remaining anchored in a game theoretical framework. Finally in the last subsection I shall stress Bacharach's conception of "salience" and "focal points". Schelling (1980 [1960]) who first brought these concepts into game theory without defining them, left the door open to subsequent interpretation and application within games. In this perspective, Bacharach purports to build a "theory of focal points".

⁹ A *valid* Theory of Games must be grounded on a *meta-theory of games* for Bacharach (1987, 1994). More specifically, a *meta-theoretic game theory* "does not regard the game directly, but regard a feature of the game theorists' theory of it, namely the 'logical closure of the information attributed to the players in this theory's assumptions" (Bacharach, 1994: 8). For instance, "it allows one to focus attention on the logical relations between what players know just before they choose and the 'normal form information' with which they are traditionally credited; more specifically on how they know just before they choose depends on what can and cannot be proved from their normal form information" (*ibidem*). Besides axioms must specify the structure of games, they enclose: "the 'set-up'" (the available actions and preferences); the players' knowledge about this and about each other; and what it is that makes actions "satisfactory" and how "satisfactoriness" relates to choice" (Bacharach, 1987: 17). In addition, according to Bacharach, in game theory "solution-concepts are established in theorems" (*ibid.*, 18). It entails that "if one rational agent described by the theory is bounds by its logic to act in a certain way (the way specified in the solution-concept), then other players also described by it are bounds by its logic to know that he will so act" (*ibidem*). Hence, a *valid* Theory of Games must specify what is rational for a player to do, given his information.

¹⁰ As Bacharach emphasizes: "in the [...] 'framing effect' literature, an experimenter studies the sign of the effect on choice of a manipulation designed to affect subjects' frames. In this part I turn to how people tend to frame the situations they face in the absence of manipulation" (Bacharach, 2001: 4). In addition, when he mentions his conception of frames he always refers to Kahneman and Tversky (Kahneman, 2000; Kahneman and Tversky, 1984; Tversky and Kahneman, 1981, 1986); even though, again, he uses framing in a different manner than they have.

¹¹ See for instance Keren (2011), for a classification of the different accounts of framing.

2.1. Framing and gaming: VFT's drawing boards.

In order for the VFT to be a *valid* theory (with respect to Bacharach's definition of a valid theory), its formalization requires three main premises:

- i. The model must specify players' frames, i.e. players' *representations* of the decision problem they face and "what determines players' frames" (Bacharach, Bernasconi, 1997:5). This requires a description of the process by which frames come to players' minds and an assertion on the structure of players' frames.
- ii. According to their frames, players must have a set of alternatives, i.e. *subjective strategies* – among which they have to make a choice. Distinguishing subjective strategies and objective strategies means that it is the players themselves who establish their own set of strategies (which may differ from the strategies considered by the game theorist in the games matrix).
- iii. Finally, principles of *equilibrium selection* and of the games' equilibrium must be involved (*Ibidem.*)

These premises allow Bacharach to dwell into a game theoretical framework¹². Nonetheless, the first two premises violate some implicit rules of standard game theory. First, in standard theory of games, the players' conception of their decision problem does not matter and even more cannot vary. What matters to resolve a game is only contained in the matrix – specifying the combination of the available actions for each player and their 'desirability'. To the contrary, in VFT, games are no longer self-contained. It is impossible to build a game without considering its context. Second, if players make their own strategies they may no longer act according to the games' rules – roughly speaking those that are defined by the game theorist¹³. In VFT players have to build their own rules: "players must derive "rules" – *formulate their problem to themselves as a certain game* – from some initial apprehension of their situation" (Bacharach, in Binmore et al., 1993: 258; my emphasis). This statement entails that: (i) who are the players matter and (ii) the way they label their strategy matter. Moreover it can imply a dissymmetry between players which go beyond, and is even more disturbing than, a simple informational dissymmetry – resolvable by Bayesian rationality, i.e. by belief updates. For instance, players can have different strategic spaces. At this point we already understand that Bacharach has to deal with a difficulty. As soon as each player builds herself her own game, Bacharach must define on what basis players will interact.

Players define the game they are going to play through a two steps process:

- i. A first *unconscious phase*: framing; "[t]he way people conceive their options is typically beyond conscious control" (Bacharach, 1991: 1)
- ii. A second *conscious phase*: the "reasoning phase" (*ibid.*, 3), in which *principles of rational choices* are operative.

Roughly speaking the object of the first stage of the VFT is to "offer a rudimentary *descriptive model* of the *process* that brings questions to the minds of the players and the consequences of this process" (Bacharach, in Binmore et al., 1993: 258; my emphasis), whereas the second stage is the *game per se*.

When "normal" agents frame a decision problem, they have *families of concepts*, which come to mind¹⁴. A family can be understood as "classifying things along a single dimension" (Bacharach,

¹² In this manner, VFT remains anchored to an axiomatic method. There is an axioms-set about the structure of the game and players' knowledge, and Bacharach gives principles of equilibrium selection entailing that the principles of rational choices are defined

¹³ To quote Hargreaves Heap and Varoufakis (2004 [1995]: 31), in standard game theory, "individuals know the rules of the game: that is, they know all the possible actions and how the actions combine to yield particular pay-offs for each player".

¹⁴ Bacharach gives variably the labels 'concept', 'attribute' or 'predicate', nonetheless the same semantic remains. I must point out that the label 'attribute' is the usual term to which psychologists refer within the framing literature,

1991: 3). Concepts are basically characteristics or ‘*properties*’¹⁵ of the objects; for example being green, red, round, etc. In these cases the families to which these properties belong are ‘color’, ‘shape’, etc.

As Bacharach builds his VFT with matching games¹⁶, i.e. games in which players have to choose objects and coordinate by choosing the same object; concepts are assimilated to characteristics, i.e. to properties of the objects. We can argue that it corresponds to a naturalistic interpretation of what things are (it is as if there was a simplistic and universal way to represent the objects). There is no semantic in the characteristic associated with the objects. For instance, players could possibly interpret or give different semantics to the colors they perceive. As a consequence, their representations would differ and accordingly, they may not think and deliberate from the same premises.

Frames must respect three main properties:

- i. A family is supposed to be *separating*, i.e. two families must be distinct. If, for example, f and f' are two families of concepts, b , a given concept, cannot belong to both f and f' (*ibid.*, 10).
- ii. The set of families, termed a “repertoire” (1993, 1997, 2006), represents a “partition” of all the concepts that describe the situation. Every concept belongs to a family (and only one family) (*ibidem*). A “generic” family¹⁷ which comes to mind with a probability of one, entails that each situation can be described. It is basically equivalent to say that every thing can be described by the concept ‘objet’, allowing players to at least choose something. As a consequence, under any circumstances, decisions problems are always describable, even if the way by which decisions problems are describable always varies (Sugden, Gold, 2006:15)
- iii. If one concept of a family comes to players’ mind the other concepts belonging to the same family too (*ibidem*). That is why Bacharach (1993, 1997) claims that concepts come to mind into “clusters”. He gives the following example: “if I notice that the blocks are square, I am likely to notice that others are triangular” (Bacharach, in Binmore et al., 1993: 259).

Accordingly, players’ repertoire represents their frames, i.e. the set of families of concepts they handle. Depending on the context some characteristics of the objects and therefore the associated families “instantiate”¹⁸. The likeliness of a characteristic – and its associate family – to come to players’ mind is formalized by a probability called the “availability” of the family (Bacharach, 1991: 17)¹⁹. This notion of availability is a kind of “propensity” to perceive something and the fact is that, for Bacharach, this propensity can be “interpreted as that of a normal person in the circumstances” (*ibidem*). Framing is a matter of salience and noticeability²⁰: “whether a normal person thinks of a predicate [i.e. a characteristic] in a situation depends on how conspicuous or noticeable is the feature of the situation which the predicate expresses” (*ibid.*, 16). Accordingly, “highly conspicuous features are more likely to come to mind” (*ibidem*). Nonetheless, availability may not only be a matter of context and environment, but may depend on players: “frames may vary [...] across players” (Bacharach, 1997: 4); and more precisely on players’ culture: “[the

whereas the others are specific to Bacharach

¹⁵ Bacharach does not explicitly refer to properties. Though, in his contributions on VFT, in practice, concepts are always equivalent to properties possessed by the objects players have to choose.

¹⁶ A matching Game is defined by Bacharach as “a pure coordination game in which there are two players with the same act-set; both get a prize if and only if both choose the same act; and the prize is the same whatever this act may be” (Bacharach, Bernasconi, 1997:2)

¹⁷ Gold and Sugden (2006: 15) refer to the term “generic frames” to describe this “generic” family.

¹⁸ Bacharach does not refer anymore to this concept after his first paper in 1991. Even though, the principle remains and is included in the concept of “availability”.

¹⁹ As stated above, in the later work on VFT, i.e. after 1991, the availability function encompasses the idea of instantiation and approximates the instantiation function.

²⁰ More specifically, in Bacharach’s models of VFT, availability is almost a matter of “visual noticeability” (1991, 1993, 1997).

availability function] is given by the culture to which the players belong” (Bacharach, 1993: 267)²¹.

Players’ set of strategies are defined from a set of options; “an option of a player is a feasible action for her described as she herself describes. It follows that a player’s options must be descriptions of possible actions which only use attributes [i.e. concepts] in *her frame*” (Bacharach, Bernasconi, 1997: 6). Again we see Bacharach’s particularity. In standard game theory there is no need to label a strategy in such a way, since the ‘desirability’ of each outcome is sufficient for players to make their decisions. More specifically, Bacharach and Bernasconi assert that “a player has one option for each attribute belonging to a family” (*ibid.*, 7). From each available concept, an act description follows (Bacharach, 1991:16). In a matching game, this conceptualization allows only two types of actions :(i) “choose/pick” the unique object which possesses the property *b*, or (ii) “choose/pick at random” one of the objects when none possesses a unique characteristic.

An ambiguity appears. Players’ strategies seem to be only subsets of their available options. They depend on the others’ sets of options and, in turn, their strategies, defined in the same way, i.e. defined within their representations. Indeed, in each game matrix, as it is framed by the players, only a subset of their available options remains. However, Bacharach does not have a clear distinction of these two concepts. This first ambiguity raises another one, more noticeable in the 1991 paper. There is no clear account in Bacharach’s models of VFT, on when the first unconscious phase ends. It seems that when players establish their set of strategies they are already deliberating as they choose their strategies “rationally” by taking into account the possible options and thereafter strategies of the others. They eliminate the options they consider to be irrelevant according to the interaction they are going to have with the others. Bacharach only specifies that at the end of this first unconscious phase the game is determined: “this phase determines a game in which they [players] have the options they conceive” (*ibid.*, 2). Hence, referring to this quotation we can extrapolate that the process which determines the game still belongs to the first phase. But in Bacharach’s VFT deliberation regards the game itself, i.e. the moment in which players resolve the game. Thus, Bacharach does not seem to distinguish the two steps to which I refer. This ambiguity will be less evident in the papers after 1991. Bacharach asserts in these other papers, that players’ options are the final strategies. As a consequence, the two steps to which I refer become less obvious. Nevertheless, the uncertainty regarding the end of the first unconscious phase remains. This problem may be considered alongside with the fact that “actions need to be individuated by the theorist in some way that is independent of agents’ descriptions” (*ibid.*, 5); and so it is for players’ strategies. We understand that in one way or another, the theorist must ‘objectivize’ the game or at least some of the fundamental aspects of the game²².

As soon as the framed game is defined, or shall we say, the ‘objectivized framed-game’, *players choose strategically* their actions between the two kinds of actions previously specified (i.e. “pick one object” or “pick at random”). Bacharach (in Bacharach, Bernasconi, 1997: 4) claims “strategies are chosen in a way which is *rational in a perfectly familiar game theoretical sense*. However the game that gets played is determined by non-rational (though not irrational) features of the players. These are players’ “frames” ” (my emphasis).

Bacharach, advocates for four principles of “rational decision”. They are not always found or equally developed in the different models of VFT, since they are linked. The first ones are the following:

- i. “A principle of coordination” or, in other words, a principle of “payoff dominance” (Harsanyi, Selten, 1988), meaning that if an option is Pareto optimal, and if “players have sufficiently strong mutual knowledge with the other of this fact” (Bacharach, 1991: 3), it is

²¹ The consequences of such a statement will be detailed in the following subsection.

²² This is tightly linked to the fact that availabilities must be independent of players. A point I shall discuss in the next subsection.

rational for them to choose this option and only this one²³. In a game with multiple equilibria, if players are rational they play their “part” “in the unique Pareto-optimal equilibrium if there is one” (Bacharach, in Binmore et al., 1993: 257).

- ii. “A principle of insufficient reason” entailing that if options are perfectly symmetrical (i.e. if none of the characteristics, concepts or families, allow players to differentiate them), it is not rational – in other words, there is no “*sufficient reason*” – to choose one of them. For instance, “if two options *a* and *a'* are alike in all relevant respects, a solution concept should not pick out one rather than the other” (Bacharach, Bernasconi, 1997: 11). In matching games these relevant aspects for Bacharach are families and frequencies (that is, the probability for families to instantiate).

The two others principles explaining and justifying the two principles of equilibrium selection above, considered to be “general feature of VFT solutions” (*ibid.*, 10), are: (i) the “principle of rarity” and (ii) the “principle of availability” (Bacharach 1993: Bacharach, Bernasconi, 1997). The former means that a player prefers to choose a rare object. By doing so, and against several identical objects, she maximizes her payoff: “*Ceteris paribus, players prefer to pick an object which is rarer*” (Bacharach, Bernasconi, 1997: 10). The latter implies that, in some circumstances, choosing a rare object can be too risky, so a player prefers a more available object: “*Ceteris paribus, a player is more inclined to pick an attribute which is more available*” (*ibid.*, 11). Bacharach and Bernasconi (*ibidem*) qualify games in which “availability preference” competes against “rarity preference” as “tradeoff games”.

To understand a tradeoff game I quote an example given by Bacharach (in Binmore et al., 1993: 269):

“[In a matching game] a grain-aware player should mark the wavy block if the probability that the other is grain-aware exceeds the relative rarity of red blocks, and should mark the red one if it falls short of it [...] [but] it may be that her partner is grain-unaware. In that case he doesn’t have marking the wavy blocks on his list of options, and there is no reason to think that he will mark it [...]. Indeed, picking a red will appear to him the best option to coordinate on. So it is riskier to mark the wavy block than to pick a red, because of the relative availability of the grain-pattern concepts”²⁴.

As emphasized in this quotation, the probability that a player perceives the grain of the object may be too low compared to the probability that she perceives the color. In this case, the player chooses the more available object, i.e. she picks at random one of the red ones. By this way, she maximizes her subjective expected utility. Here again we understand that players’ strategies, which are presently defined as their options, are only a subset of their respective available “act-description of actions”.

Conceptually, considering these principles of rational decision, i.e. of equilibrium selection, the game is resolved if each player finally chooses an option. Bacharach (1991: 20) asserts that a player chooses this option “by arguing validity from her ‘initial knowledge’ [...] that she should choose it”. He defines player’s initial knowledge as follows: “[h]er initial knowledge is the knowledge that is ascribed to her by the definition of the MG [Matching Game]” (*ibid.*, 20). Therefore, players’

²³ Bacharach admits that this principle can be controversial, but he justified its application by improvements in game theory which lead game theorists to recognize the principle of coordination as a rational way to play in games. He refers especially to Aumann and Sorin (1989) and to Anderlini (1990). In VFT, the accuracy of this principle is mainly based on the epistemic assumptions (as players must have “a sufficient strong mutual knowledge” on the fact that a Pareto optimal equilibrium exists). This fact explains why Bacharach is concerned by a rigorous definition of players’ knowledge. Besides, we must mention Colman’s attempt (2003 : 144-145) to prove the relevance of this concept ; objecting to Gilbert (1989), and referring to Crawford, Haller (1990), Farrell (1987, 1988), Gauthier (1975), Janssen (2001b), Lewis (1969) and Sugden (1995, 2000) .

²⁴ We understand in this quotation the slight methodological changes to which I was referring above, occurring between the 1991 paper and this one. Options are presently players’ strategies. But again, as I previously asserted, the doubts concerning the end of the first unconscious phase are still true.

initial knowledge does not merely result from what they perceive and what they are aware of, but is linked to the structure of game (even if this structure emanates from players' perceptions). It encompasses the knowledge players have about the others (their strategies and hence the 'desirability' of each outcome), in other words, players know the rules of the game. This statement induces some difficulties and ambiguities in VFT (which I shall emphasize in the next subsection). Formally, a game solution is a pair of subjective best reply, i.e. a pair of "decision functions" (Bacharach, in Binmore et al., 1993; Bacharach, Bernasconi, 1997)²⁵. This principle is similar with standard non-cooperative game theory. In the latter, a player is rational if she maximizes her subjective expected utility, which, in other words, means that she adopts a principle of best reply. And like every player is defined in the same way and is rational in the same way, a solution is necessarily a pair of subjective best reply. Accordingly, what does Bacharach qualifies as a solution involves exactly the same principles as standard non-cooperative game theory. Indeed, in the VFT, each player maximizes her subjective expected utility (see, e.g. Bacharach, 1993²⁶).

2.2. The "status" of the game: From players' intersubjective to objective representations?

Incorporating players' subjective representations of their decisional problem begs a question: how can a common game be defined among players? Bacharach must justify in some way or another a common structure within which players will be able to play. Indeed, as stated by Scazzieri (in Gavalotti et al., 2008: 187): "a critical feature of rationality assumptions in economic theory is their association with the way in which reasoning and decisions by *heterogeneous and independent individuals* (or collective agents) may be made *consistent* with one another within a *congruent structure*" (my emphasis). Because players' representation of a common decision problem may be potentially dissimilar, Bacharach needs to warrant a "congruent structure".

To understand how possibly different subjective representations can lead to a matrix shared by every player, we must first understand the status of a game in VFT. Analytically, games matrix in VFT corresponds to the joint probability for players to choose the same object. Bacharach (in Binmore et al., 1993: 256) emphasizes that standard non-cooperative game theory "misspecifies the payoffs matrices". This statement implies that the purpose of the VFT is a *respecification* of standard coordination games²⁷. Therefore, games in which there are multiple equilibria become games in which only one Pareto optimal equilibrium remains. In this manner, Bacharach ascertains that players coordinate on the Pareto optimal equilibrium. Nevertheless it is the theorist who necessarily heads this process. It is contradictory to the fact that a game is an abstraction process carried out by the players themselves. In fact, "[a]ccording to the way that normal agents tend to 'cut-up the world' as they come to terms with a problem, this problem gives rise to payoff-matrices that differ widely in terms of the number of options and the pattern of payoffs: the payoff *structure* generated by a problem is concept-dependent" (Bacharach, 1991: 3). It does not explain how players come to a *common respecification* or whether they perceive the same joint probability. So how do players recognize the game from the moment that it is no longer an objective game but a framed game in which each player has the options she herself conceives? How can one be sure that players play the same game and have a minimal common background, from which they will be able to coordinate? This difficulty requires controversial assumptions, and subsequently induces several methodological changes.

²⁵ In the 1993 paper, an 'indicator' leads a player with a given repertoire to an action which is the best reply to the other players' indicator defined in the same way. However, his concept is not present in this other papers.

²⁶ This is not explicitly, i.e. formally, stated in 1991; however this is the case in 1993 and then in 1997.

²⁷ Bacharach (in Binmore et al., 1993: 257) insists on the fact that "this respecification is the key to the game theoretic argument I shall give for the intuitive solutions" because "given these matrices [after the respecification], we have only to invoke the so-called Principle of Coordination" (*ibidem*).

In his first contribution on VFT in 1991, Bacharach argues that among players, and especially ‘normal’ players, there is a tendency to perceive the situation in the same way, and to understand this fact. This assertion is explained by the combination of different elements. First, players’ frames are structured in the same way, i.e. their frames are constituted by sets of families coming to their mind. Second, “the propensity [to perceive such or such thing] is interpreted as that of a normal person in the circumstances” (*ibid.*,17). This statement implies that each player has *in fine* the same repertoire because each player is normal. Third, knowledge in VFT is an “occurrent knowledge”²⁸. Some facts are considered to be N-evident for Bacharach as soon as they are observable: “one condition for p to be N-evident in S [e.g. in a given situation, say S] is that any normal *thinks of* p in S ; another, that in S any normal who thinks of p is able to confirm that p (e.g. through observational evidence available to her normal sense) – that is, p is ‘N-confirmed’ in S ” (*ibid.*, 16). Basically, entailing that normal agents think p means that they are aware of p because p is a proposition that has come to their mind by the framing process described above. Then, “if (in a situation S) p is N-evident then, if p , a normal knows that a normal knows that ... that p . Now suppose that (in S) p is ‘N-evident *occ F*’ [e.g. N-evident conditioned by the occurrence of the family F to which the proposition p belongs], in the sense that if p is true then any normal player knows *occ F* that p . Then we get an analogous bonus of layers of iterated knowledge: it can be shown that if p is true then a normal knows *occ F* that a normal knows *occ F* that ... that p . If it is also an axiom that normal recognize each other in S , this strengthens to: if P_1 [player 1] and P_2 [player 2] are normal and p is N-evident *occ F* then, if p , each knows *occ F* that ... that p ; that is, if p is N-evident *occ F* and true, there is mutual knowledge *occ F* between P_1 and P_2 ” (*ibid.*, 17-18). That is why, *in fine*, the representation of the situation is not only common among all the players who face this situation (because their frames are equivalent) but it is mutual knowledge too (*ibid.*, 3). Now, because players know they can act on their frames, if a concept, say k , instantiates, players know they can act on k and accordingly possibly choose the object possessing the characteristic k accordingly. As a consequence, because each player has equivalent frames and knows she can act on each concept of her frames: at the end each player possesses the same set of act-descriptions, and this fact is of mutual knowledge. In this manner, Bacharach (*ibid.*, 3) asserts that he “provide[s] an endogenization of mutual knowledge between the players of the game they are playing, described in a certain way”. Finally Bacharach adds that some fundamental characteristics of the game are of mutual knowledge²⁹. These characteristics merely ensure that all players mutually know the rule of the game, and play according to these rules. What follows is the fact that players are going to play the same game.

To postulate that everyone perceives the same thing from the same situation, and mutually knows this fact is very restrictive. It is probably one of the most controversial assumptions made by Bacharach. It may be plausible in very specific contexts. For instance it can be true if players are closed enough, in the sense that they regularly face the situation together and interact. Knowing each other quite well implies that, in a way, individuals know how the other perceives the situation and thinks. In this case and only in this case this assumption can make sense. However in other cases it is even more overconfident than standard non-cooperative game theory. Within standard game theory, a payoff matrix pre-exists before the game, and can be considered as an objective material on which players interact strategically. It is even more restrictive in Bacharach’s model to

²⁸ Asserting that players knowledge is “occurrent” means that a player knows something if it has come to her mind, i.e. if this thing occurs (Bacharach, 1991: 15; Bacaharch, in Binmore et al., 1993: 257).

²⁹ These characteristics are the following: (i) the “admissibility condition”, i.e. if there is only one object which possesses a given characteristic a player must choose that object; (ii) the “preference condition” i.e. players attach utility 1 to the fact that they choose the same object or 0 otherwise; (iii) the number of objects which possess a given characteristic, say, b ; (iv) the fact that a player can pick a non-unique object at random; and (v) the fact that each player is rational in the sense in which Bacharach defines rationality in this paper (i.e. *via* the different principles of rational choice defined above).

presuppose that each player, as a normal agent, comes to build herself, and *independently*, a game which is common to everyone. We could almost advocate that this model does not fit standard non-cooperative game theory: the uncertainty is quasi null. Moreover, Bacharach never refers to beliefs, as if beliefs were not necessary in his model. This having been said, we are no longer in a strategic context. We are outside a game theoretic framework.

That is why in Bacharach's second paper on VFT in 1993, an important methodological change occurs. Bacharach no longer makes reference to 'shared' perceptions. Hence, the role of beliefs becomes predominant: "[t]his model gives us probabilities for players' having various belief-spaces. It allows us, too, to say something about the beliefs which a player with one belief-space has about the belief-spaces of others; and *this is vital if we are to have a theory of encounters between such players that's worth calling a theory of game between them*" (Bacharach, in Binmore et al., 1993: 258; my emphasis). It seems that Bacharach was aware of the fact that game theorists could object to the precedent model, saying that it was not game theoretic. In the 1993 paper, players have to form beliefs about their available choices – according to their frames –, others' frames and accordingly others' subsequent choices and acts. "If player has repertoire r , then for each a [a is a given action] in $A(r)$ [$A(r)$ is the act-description space induced by the repertoire r] she believes that she can a [...], and for each a not on $A(r)$ she has no belief as to whether she can a " (*ibid.*, 260). It means that she cannot believe that she is able to handle an action which is not based on her repertoire (*ibid.*, 260)³⁰. In addition each player has a subjective probability for every sub-repertoire of her own repertoire. This subjective probability is interpreted as "the subjective probability that a player having repertoire r assigns to the other player's having r' " (*ibid.*, 261). A player cannot attribute to others the same repertoire as her own: " P_i implicitly assumes that P_j 's repertoire ($j \neq i$) is some sub-repertoire of hers; that is, if she has r , she implicitly assigns probability 0 to his having occur to him any family not in r " (*ibidem*).

Hence, because players' perceptions are not necessarily shared, it may be impossible to derive a common game from players' subjective representations of the decision problem they face. Each player herself computes her own game, making assumptions about the other players' frames and choices. Therefore, even the game is subjective. Subsequently, to build a game Bacharach makes a restrictive statement: players must have the same set of act-description of action. Although this postulate justifies the structure of the game, and makes sense because players believe that others' set of act-descriptions is necessarily a subset of their own (and accordingly their own set of strategies is symmetrical vis-à-vis the set of strategies they ascribe to others); it does not justify the fact that two or more players agree on the game they are going to play. Nothing ensures that players can independently build a common game. This is indeed the major limit faced by Bacharach's attempt to enrich standard non-cooperative theory by players' subjective frames. The present risk is that no common game can be specified.

As a consequence, a convergence among players' perceptions is required. Two options are available: (i) asserting a universal way to represent situations, and (ii) asserting that representations are socially constructed but, in this case, repeated games are necessary. Subsequently, the only way to escape from this pitfall is for Bacharach to postulate that there is a universal representation of a given situation, and that everybody can have access to a part this universal representation. This was implicitly postulated in 1991 and it finally remains in his later work in 1993, 1997, and 2006, even

³⁰ Bacharach's conception of knowledge evolves to a diminished version. Indeed, players *believe* that they can operate such or such action depending on their frames whereas previously they *knew* that they were able to choose this or that action. This methodological change can be considered alongside the ambiguity about the switchover from the set of act-descriptions to the strategies, I have highlighted previously. Asserting that players have to form beliefs about their options means that they have to take into account the others' set of act-descriptions of actions to select their relevant options within their own. Accordingly, Bacharach gives an explanation of the difference between act-description of actions and strategies.

if the methodology and the formalism are sensibly different. This postulate implies that every situation, every characteristic of the world, every object, is objectively describable and perceivable. However, Bacharach and Bernasconi (1997: 7) claim that in reality: “there certainly are attributes, such as colors, comparative sizes, and simple geometric shapes, whose saliencies are universal constants (Treisman, 1986). But for others, saliencies are culture-dependent”. Analytically, this cultural determinism is discarded. Bacharach dismisses all social (and cultural) determinants in players’ interaction. Players do not use their interactions to build themselves the games they are to play. To some extent, games are merely *natural abstract representations* partially grasping the universal representation.

In addition, Bacharach is formally constrained by the fact that (i) allowing as many conceptions or perceptions of the situation as there are possible (natural) conceptual representations, and (ii) “different possible descriptions may be strategically inequivalent” (Bacharach, 1997: 5), which entails that the game may not be defined. Thus, Bacharach has to postulate that *in fine* families’ probability of occurrence is independent from the players (1993, 1997). More precisely, these probabilities are independent both across families and across players. Subsequently, frames are only explained by the context even if he recognizes that it is highly simplified. We clearly understand by saying that, that the naturalistic interpretation of Bacharach conception of framing is reinforced. In other words, as the theorist, Bacharach endows players with given objective frames *a priori* – because the probability distribution over families availability is objective and stated *a priori* – and he specifies the rules of the game and establishes the epistemic requirements of the models. He necessarily abstracts the game from players’ subjective perceptions, and portrays the game as an objective fact to ensure a coordination device. This is the only way to define a game, according to Bacharach’s methodology. Indeed, it is the only way to conciliate the players’ subjective representation, his conception of what a game is in his VFT and the fact that players must play together in a common game.

The purpose of the next subsection is to highlight that this conception of games induces a problem of consistency in Bacharach’s account of salience and poses a difficulty to disentangle the nature of salience.

2.3. Games’ solution and focal points: a ‘naturalistic’ or a ‘community-based’ approach?

To “outline a *rigorous theory of salience*” (Bacharach, Bernasconi, 1997: 2), Bacharach (in Binmore et al., 1993: 256) purports “to solve the unsolved problem of how to give a game-theoretic rationale for choosing the salient”. For him, “a satisfactory theory should show whether these choices [i.e. choosing the salient options] are indeed rational and, if so, why and what their rationality has to do with salience” (Bacharach, 1991: 5). Accordingly, he claims, “I will show how the framework of MGs [i.e. Matching Games] makes possible precise definitions of salience concepts – even if it is not easy to say which is best. I shall show, secondly, that we can say precisely what the connection is, in MGs, between choosing the salient (in a defined sense) and choosing rationally” (*ibid.*, 34)³¹. In Bacharach’s account of a theory of focal point, being a salient option should not be “reason-giving”: “in order to explain why rational players choose salient options, we need never invoke the salience of an option” (*ibidem*). To the contrary, it should be the result of a rational choice. That is why VFT “derives focal-point play from an explicit model of salience and from well-defined rationality postulates, and it also makes predictions where the traditional theory is silent” (Bacharach, Bernasconi, 1997: 37).

I will attempt to show how Bacharach unifies (i) his conception of a game theoretic solution, (ii) the equilibrium selection understood in terms of *defined rational choice*, and (iii) the fact that

³¹ It is interesting to note that, for Bacharach (1991: 34), the definition of salience depends on the theoretical frameworks in which this notion is applied: “different notions may be defensible in different theoretical contexts”.

games are the output of a subjective and unconscious process of players' abstraction of a strategic decision problem. Indeed, in Bacharach's conception of game theory, a solution must be strictly defined; it is the core of the axiomatic approach of game theory³². In addition, "a determinate theory of rational play known by a rational player indicates to the player the right action to choose, that is, brings such a player to believe that a certain action is the right one" (Bacharach, in Binmore et al., 1993: 264).

To define salience, Bacharach starts with Lewis' definition (1968), i.e. salient options possess two characteristics: (i) 'conspicuousness' or 'noticeability' and (ii) 'uniqueness'. According to Bacharach (in Binmore et al., 1993: 270) these two dimensions must be distinguished: "[conspicuousness and uniqueness] are logically independent, although they frequently appear in amalgam in discussion of salience". He does not assert that they are totally independent but he insists on the fact that they involve different mechanisms in individuals' process of reasoning³³. In addition, because the definition of salience is "theory-specific" for Bacharach, he has to make the link with framing. For instance, "a frame is said to be salient if it has a strong tendency to be operative; some frames are more salient than others; and the salience of a given frame depends on the context. The salience of a frame depends on the salience (similarly defined) of its constituent concepts, and in particular of any constituent classifiers: for example, the colour-shape frame is salient if both the colour and the shape classifiers are" (Bacharach, 2001a: 5). Besides he affirms "frames can have more or less power or potency to influence decision" (*ibid.*, 5-6). The last quotation is explained by the fact that in VFT, players must think of salience strategically: "the essential depth assumption of VFT is that players think about salience strategically: *i* asks herself how likely it is that *j* has noticed what she has and so has the same options as her" (Bacharach, 1997: 39). As a consequence, two possible interpretations of focal points and salience could compete: (i) a "naturalistic" and (ii) a "community-based". The former implies that salience ensues from "objective and natural properties of some entities (events, strategies, outcomes)" (Hédoin, 2012: 2; referring to Lewis' naturalist interpretation of salience). In this account salience is an objective fact, directly observable in the environment, and above all, understandable by everyone in the same way³⁴. Conversely, social institutions explain the latter. Individuals belonging to the same community inherit the awareness of some institutional facts from this community. Furthermore, they understand and interpret these institutional facts in the same way, and know that³⁵, but they must be able to recognize each other as member of the same community³⁶. The two competing accounts of salience in VFT are stressed by a methodological change occurring between 1991 and 1993, 1997, 2000, and 2006.

In 1991 Bacharach claims that the two dimensions of salience in Lewis's definition do not belong to the same phases in the players' reasoning process. That is why "to explain how salience brings

³² In particular, "the connection between being an equilibrium and being a solution of a VU-game [i.e. games within the VFT] is given by applying to these games the indirect argument – the general argument that being an equilibrium is a necessary condition for being a solution if there is a determinate theory of rational play (von Neumann and Morgenstern 1944; Luce and Raiffa 1957; and Bacharach 1987)" (Bacharach, 1993:264). If this statement is not respected VFT fails to be a game theoretic model. Nevertheless, as players are supposed to build the games they are going to play themselves, the existence of equilibria and solutions stated *a priori* may be hard to conciliate.

³³ "[A]lthough the two aspects of salience must be kept separate in explaining salient choice, this does not mean that they're causally independent of each other. Probably they are not: for if a property is uniquely instantiated this, *ceteris paribus*, makes it no more conspicuous" (Bacharach, 1991: 36).

³⁴ Referring to Postema (2008: 43-44), Hédoin (2012: 18-19) claims that naturalistically, "the salience of an entity [...] is taken to be a brute, objective and explicit fact. Salience is then a characteristic constitutive of an entity which can be recognized "straightforwardly" by agents"

³⁵ "[B]ecause individuals in the same community share the same institutional heritage, they have confidence in the fact that they interpret the same institutional fact in similar fashion. They have a common understanding of many situations and they are used to rely on the same institutional signals to make sense of what is happening" (*ibid.*, 16).

³⁶ "[C]ommunity membership is more or less a public event because there are many signs that may allow identification of the community to which one belongs" (*ibidem*).

about choice it is therefore necessary to decompose it and to explain the operation of the separate components in these two quite different mechanisms” (Bacharach, 1991: 34). Conspicuousness relies on the unconscious phase³⁷ whereas uniqueness results from the conscious phase. More precisely, “a property of *bs* is conspicuous just if *a* predicate denoting it is highly available” (*ibid.*, 34). This is what Bacharach calls the “salience-1” and by definition “salience-1” belongs to the unconscious phase. “Salience-2” implies that “*b* is ‘salient-2’ for P_i if it’s the only *b* such that for some highly available family *F* she knows that it’s the odd one out in terms of *F*” (*ibid.*, 36). In other words “salience-2” “raises the expected utility of the pair in which both choose the salient option, and so give a game-theoretic reason for choosing it (expressed in the principle of coordination)” (*ibid.*, 35). However Bacharach slips this in other subsequent versions. He no longer dissociates the two dimensions of salience into an unconscious and a conscious phase. After 1991, salience relies entirely on the availability of frames. The two dimensions of salience are after all linked to the properties or attributes of the objects. And the fact is that these two dimensions may be linked to different objects. If this is the case, Bacharach identifies a “trade-off” game³⁸. In this type of games, a “trade-off theorem” allows players to make a rational choice, thanks to the “principle of availability” and “principle of rarity” (cf. 2.1). As a consequence Bacharach distinguishes a “primary salience” and a “secondary salience”. For him, the former is “the traditional notion” (Bacharach, Bernasconi, 1997: 38) – i.e. noticeability *and* conspicuousness of a property; whereas the latter entails that an option is of “secondary salience for a player if she thinks it has primary salience for her coplayer” (*ibidem*)³⁹. Although Bacharach gives two dissimilar justifications for the difference he makes between conspicuousness and uniqueness, the underlying idea remains: *one dimension of salience is strategic* (which merely means that one dimension relies on the unconscious phase and the other is explained by the deliberative phase, i.e. the strategic phase).

How can the possible different competing interpretations of salience be disentangled in Bacharach’s VFT? In fact, some relevant aspects of salience seem to rely on the object’s natural properties whereas some other aspects rely more on the players who are interacting.

When Scazzieri refers to Bacharach’s account of framing and salience, he emphasizes: “framing shows the *interplay between structural principles and evolutionary (historical) principles*” (Scazzieri, in Gavalotti et al., 2008; 197; my emphasis)⁴⁰. While these “natural associations” between frames and context are partly explained by history, e.g. culture, it means that they are socially based; and accordingly, we can interpret salience as “community-based”. If players’ subjective frames are partly explained by precedents, players can ground their reasoning on what they learned about coordination devices inherited from this culture, and accordingly the way others generally perceive the coordination context they face. Since players think of salience strategically, it means that they think about what is salient for others; they have to refer to institutional dimensions of the culture in which the other is embedded; hence they use the knowledge their culture gives them. Indeed, players cannot think of others’ perceptions if they are not from their own background (Bacharach, 1991, 1993, 2001; Bacharach, Bernasconi, 1997), which implies, in a way, that each player’s perception must be embedded in a common conceptual framework, i.e. in a common cultural legacy. Moreover, all of the VFT models incorporate language. Players must share a common language⁴¹ to play a game. These facts reinforce the suggestion that salience may be

³⁷ “[C]onspicuousness is played out by in the invol[un]tary description phase of the matching problem, where it brings it about that *O*ing a *b* [*b* being an object; *O*ing a *b* means choosing this object] having the conspicuous property is an option” (Bacharach, 1991: 35).

³⁸ In a trade-off game, “the two aspects of salience in Lewis’s definition – obviousness and uniqueness – are possessed by different objects” (Bacharach, 1993: 270).

³⁹ Bacharach refers to Metha, Starmer and Sugden (1994) who propose this distinction.

⁴⁰ This account suggests that a meta-theory of framing in Bacharach’s VFT is required to explain why frames are structured one way or another. However nothing is mentioned concerning the ontology of frames in Bacharach’s VFT.

⁴¹ Until now, I did not mention the use of *a* language in VFT models for a few reasons which I shall now explain. On

“community-based”.

Nevertheless, for Bacharach nothing warrants that salience is shared among players: “[t]he VFT assumptions that saliciencies are shared and that complete-frame players accurately perceive this are inessential simplifications [...] Saliencies may not be shared [...] and whether or not they are, players may assess them with bias” (Bacharach, Bernasconi, 1997: 38-39). This postulate contradicts with the “community-based” account in which the property of a focal point is to induce that every player expects that everyone will conform to focal points, and that this statement is of common knowledge (Hédoin, 2012: 2). The property of focal point is almost to generate symmetric reasoning among players. Indeed, “an event will be salient if i) it is public or can be expected to be so (that is, everyone can confidently expect that everyone else has observed the event) and ii) every one infers the same conclusion from it (and knows this fact)” (*ibid.*, 7).

Because players are not necessarily symmetric reasoners⁴² in VFT, with respect to the facts that (i) players may not have the same subjective representations (and reasoning is comprised within these representations), and (ii) the focal point may not be shared; Bacharach’s account of salience and focal point do not seem to fit with the “community-based” account. As stated above, Bacharach insures symmetry among players just by postulating that *in fine* players possess a symmetrical set of strategies but this assumption is an *ad hoc* one. This symmetry is not explained by the existence of historical coordination devices, but is *built* by Bacharach as a game theorist. This methodology is even more disturbing if we consider the following quotation: “the outcome of a given interactive problem situation can vary dramatically with the conceptual resources of the players. We have observed that there are two levels of such variation: the stochastic variations of the “coming to mind” process parameterized by the availability function V, and the possibility of alternative values of the parameters of V itself. The latter variation models effects of culture. VU-games may therefore be seen as a tool for giving precise expression to the emerging view (e.g. see Kreps 1990) that the rational solutions of some games depend on the culture of the players” (Bacharach, in Binmore et al., 1993: 271); in which again, we understand that there is space for a “community-based” account. Yet once more, Bacharach’s methodology to lay a game, i.e. through a process of objectivization of games’ matrix, which he imposes as a theorist, downplays the role of social determinants – i.e. cultural determinants – in the players’ capacity to coordinate. Frames are *in fine* explained by a naturalistic interpretation of the context. By doing that, the resolution of a game does not rely on the discovering process in which heterogeneous agents, who have different subjective representations of the strategic decision problem they face, learn how to interact, how the other perceive the situation and think, by using social and cultural determinants. In Scazzieri’s words, players do not “generate some degree of congruence [...] identify common co-ordination images, and congruent structure often emerges from tacit, and agents’ active participation ‘in the creation of traditions’” (Scazzieri, in Gavalotti et al., 2008: 187; quoting Schelling, 1960: 106).

In summary, I suggest that the numerous variations of Bacharach’s VFT models, described in this subsection, translate Bacharach’s attempt to make his theory both a “theory of rational play in games” and a more realistic one, without endangering its validity. His ambition is almost to emphasize how framing is a promising framework for game theory. Indeed, he argues that VFT

one hand language is just one more conceptual tool in Bacharach’s formalism. In my opinion, it was not relevant to mention language in the discussion until this point. Its use is only explained by the fact, that players’ strategies are descriptions of action, so they necessarily require a language. On the other hand, Bacharach *a priori* postulates a given language. Players do not create their own language to coordinate. In other words and according to Scazzieri, this is not an emergent “congruent structure” created through players’ interactions. There is neither an ontology of language nor a theory of language in Bacharach’s VFT.

⁴² Even if Bacharach postulates that there is only one way to act rationally, since he identifies principles of rational decision, as soon as he claims that deliberation (which is assimilated to rationality in his account) logically derives from perceptions and that perceptions are subjective, players can have different perceptions. As they deliberate onto their perceptions they may not infer the same conclusions even if they are rational in the same way.

belongs to “a much larger research programme for game theory: the study of the perception by players of the problem that confronts them and the condition in which this issue in games of varying form” (Bacharach, 1991: 38). Even if in his first paper on VFT Bacharach explains matrix games and endogenizes conditions of mutual knowledge (of the shared matrix, and of the rules of the game), his methodological choice for doing so, I have argued, begs questioning. And the fact that Bacharach drops out the methodology which prevails in 1991, yields to numerous difficulties to justify the existence of games and focal points – as coordination devices. For instance, Bacharach recognizes that he has to make some simplifications, which seem quite contradictory to his initial purpose.

Generally, Bacharach’s VFT sheds lights on the limits of explaining strategic interactions with social and subjective dimensions within a standard game theoretical framework. Somehow, like in standard game theory, objective structure or objective determinants in players’ strategic interaction must be postulated. That is why Bacharach is entangled in two different viewpoints. He takes on the position of both a player and of a theorist at the same time. It seems that he looks at frames and subsequently, framed games, from both angles. Nevertheless, as a theorist he must require to *ad hoc* assumptions and requires to statements which contradict with his objectives.

3. Bacharach’s theory of “Team Reasoning” and cooperation

In order to show how and why it is rational to cooperate analytically, Bacharach constructs cooperation as a different mode of reasoning than the standard one (Bacharach, 1995, 1997, 1999, 2006). Indeed, in the theory of Team Reasoning (TR), the nature of choices changes when players cooperate. More specifically, players who cooperate are “profile directed reasoners” instead of “means-end reasoners”, i.e. “best-reply reasoners”. Hence, Bacharach’s main aim is to justify why players switch from one mode of reasoning to another. I shall stress how Bacharach seems to struggle with variations on his TR models between a naturalistic and an interactional explanation of cooperative behaviors. As was the case for VFT, Bacharach begins by proposing an approach which gives quite a revolutionary impression with respect to standard game theory. He first gives an explanation of cooperation that seems quite counterintuitive in a game theoretical framework, and again, like in VFT he progressively makes conceptual and methodological changes to his models to be closer to a game theoretical framework.

In this section I shall highlight respectively (i) the common determinants and the progressive evolutions of TR in Bacharach’s different contributions on this topic, (ii) a dualism between a naturalistic or interactional explanation of cooperative behaviors, and (iii) the concept of salience. The latter point emphasizes the constraints imposed by standard non-cooperative game theory regarding Bacharach’s account of TR. Bacharach constantly tries to endogenize the probability that players, in a given interactional context, adopt the TR, and then eventually cooperate. This endogenization problem is a matter of framing; as a consequence, it is tightly linked to salience. In order to underscore how Bacharach’s conceptual and methodological choices raise numerous difficulties especially regarding the endogenization problem, I shall compare his account of TR to that of Sugden (1993, 2000, 2005, 2007). Sugden adopts a totally different methodological position. Therefore his theory of TR does not display these difficulties.

3.1. Drawing boards and evolutions

Regardless of the model of TR proposed by Bacharach, as soon as a player team reasons she

is “guided by the team’s objective” (Bacharach, 1999: 118). Team reasoning is a “profile directed reasoning”. It implies that team reasoners “first identify a profile picked out by a group choice function, then identify [their] part in it, then do that” (Bacharach, 1995: 3). The group choice function, which is a collective “profile of actions”, stems from a collective utility function. It is the sum of individual utilities and a Paretian ranking (Bacharach, 1999: 120) of the “utils” – or payoffs – (i.e. numeric values) in the objective game matrix. Basically, the same principles prevail for Sugden. Team reasoning always requires two phases – or in Bacharach’s words, two “activities”: an “evaluation activity” and a “selection activity” (Bacharach, 1997: 12). Hence, a player who team reasons first computes the best profile of actions for the team – regarding the collective utility function – then chooses the “component” of this profile “under her control” (Bacharach, 1999: 120), and finally performs its component. Sugden (1993, 2000, 2005, 2007) adopts the same viewpoint. For instance, a player who team reasons “considers which *combination* of actions by members of the team would best promote the team’s objective, and then performs his part of that combination” (Sugden, in Gold, 2005: 183).

A pervasive characteristic of team reasoning in Bacharach’s models is the fact that team reasoners, to some extent, remain “strategic thinkers”; they have to form beliefs on other players’ possible behaviors. They do not cooperate systematically; it depends on the probability that others cooperate. I shall underline how this statement induces a certain incoherence in Bacharach’s models. Conversely, in Sugden’s account of TR if players team reason, they are necessarily cooperative players. Sugden constructs the TR as a *valid* mode of reasoning. As soon as players *accept its validity*, they adopt the mode of reasoning dictated by “the rationality of the joint action of the team” (*ibidem*)⁴³, which conditions the mode of reasoning described above. In fact, because TR does not impose unconditional behaviors (this fact is true for both Bacharach and Sugden), in order for players to team reason, and accordingly to accept its validity (in a game theoretical framework), they must have confidence that others players team reason too. This confidence relies on a “condition of assurance” (Sugden, in Gold, 2005: 193)⁴⁴ insuring that each member adopts team reasoning, and therefore is cooperative. In this manner Sugden ensures that team reasoning is “an internally coherent mode of reasoning” (*ibid.*, 194).

Although Bacharach is consistent throughout his work on the consequences for players’ reasoning to adopt TR, he has made slight modifications to explain the conditions enhancing TR, between 1995 and the later models (1997, 1999, 2006). The paper “Cooperating without communicating” in 1995, draws the underpinning statements of TR, and some intuitions of the underlying ideas which will be clearly developed in his later work, even if cooperation does not rely on framing. It is my opinion that his 1995 paper is particularly relevant because it underlines how some of the conceptual and methodological choices, which are problematic in the model proposed in 1995, will be pervasive throughout Bacharach’s models of TR, even with different formalisms. Moreover this paper even more than others, clearly exhibits how Bacharach’s account of TR is hardly compatible with a standard game theoretic framework (due to the concepts he applies, and contrary to Sugden). I shall stress that this problem of compatibility is mainly explained by the fact that Bacharach’s wants after all to justify and eventually explain that *realistic* players, in some strategic contexts, naturally tend to team reason. Sugden does not follow the same logic and therefore does not encounter Bacharach’s difficulties. He first and foremost aims to demonstrate that TR is a valid “formal representation” of a success-promoting mode of reasoning in a game theoretical context, an “internally consistent analysis of a mode of reasoning” (Sugden, in Gold, 2005: 185). He mainly offers a “conceptual analysis” (*ibidem*), without value judgments about its external validity or its realism. To propose a realistic collective mode of reasoning, Bacharach

⁴³ Which is reaching the highest payoff for the team.

⁴⁴ This principle of “assurance” is defined in the following quotation: “team reasoning does not generate reasons for choice unless each member of a team has reason to believe that there is common reason to believe that there is a common reason to believe that each member of the team endorses and acts on team reasoning. This is a condition of *assurance*.” (Sugden in Gold, 2005: 193).

purports to draw TR on framing after 1995. Subsequently, his TR adheres to the same logic as VFT, i.e. showing how players' subjective representations induce various possible games, and then, possibly, various modes of reasoning, i.e. either individualistic reasoning or team reasoning. However, in TR, players' frames concern the way they perceive themselves rather than the context *stricto sensu* (even if they are linked). Nevertheless, like in VFT, subjective representations of players may induce a *respecification* of standard games. Depending on the way players perceive themselves, they *respecify* the game matrix.

In the 1995 model, Bacharach formalizes two types of players: (i) the “fellow member reasoners”, i.e. team reasoners or cooperative reasoners, and (ii) the “best reply reasoners”, i.e. standard individualistic players. This ability to reason cooperatively is explained by “natural types”. Someone who is a “type T” is a “fellow member reasoner” and has a *natural tendency* to cooperate with other “type T” players, and someone who is “type T” is a *natural* best reply reasoner. Nonetheless, a fellow member reasoner cooperates if she is “sufficiently” sure that her co-player reasons in the same way and cooperates (Bacharach, 1995: 1). Therefore, players must have the ability to recognize each other; they must be “good enough at recognizing each other” (*ibid.*, 2). Thus, a “fellow member reasoner” has the ability to switch from cooperative to standard choices: “a cooperative reasoner is a player who uses cooperative reasoning some of the times: she uses cooperative reasoning against putative cooperative reasoner, but she BR-reasons [i.e., best reply-reasons] against putative BR-reasoner” (*ibid.*, 8). When a fellow member reasoner faces a player who adopts a BR-reasoning he is “rationally forced to play a best reply to her expectations” (*ibidem*). The last two points suggest that Bacharach requires to an *ad hoc* assumption to countervail his conceptual choice, i.e. to justify that “type T” players do not invariably cooperate (and behave against their natural tendency) and accordingly that his model remains game theoretic.

In this model, Nature moves before the game, and twice. The first move induces that players observe their own type (they can be of type T or T'). The second move reveals to players the others' “demeanour”. There is an objective joint probability distribution over players' types and players' “demeanour” (*ibid.*, 12). Then players are able to choose *strategically* to cooperate or not.

This conceptual difficulty partially explains that in 1997 Bacharach no longer resorts to “natural types” to explain cooperation but to frames. In fact, the link between VFT and TR is explicitly made, and well developed, in this 1997 unpublished paper and a “variable frame model” explains “We-reasoning”, i.e. team reasoning (Bacharach, 1997: 2). Nonetheless, Bacharach still refers to the concept of “mutual recognition”, i.e. “the psychological capacity of persons with a certain disposition – e.g. to cooperate in some sense – recognize each other” (*ibid.*, 2-3). In this model, instead of applying frames to context description, frames concern how players perceive themselves and these different perceptions affect their respective reasoning. Bacharach presents three types of frames: (i) “I frames”, (ii) “W frames”, and (iii) “S frames”. I and W frames are qualified by Bacharach as “basic” frames entailing a “simplistic” way of reasoning (*ibid.*, 2). They imply “I/he” concepts and “we” concepts respectively, and induce reasoning of the sort “what shall I do?” and “what shall we do?” (*ibid.*, 5). Therefore, players in “I frames” are individualistic and instrumental reasoners, i.e. they use a “BR-reasoning”, whereas players in “we frames” are profile-directed reasoners. When players are in “we frames”, they *respecify* the objective game, i.e. the initial matrix, in a “group payoff matrix” which is basically a formal representation of the collective “utils” (*ibid.*, 11). The fact that “we frames” induce simplistic modes of reasoning raises a problem in certain types of games: “if the objective of the W player [the player in a “we frame”] is to achieve the group best, she does not necessarily achieve her objective by playing A [e.g. cooperate]. [...] if her coplayer chooses B [e.g. defect], she will do better not only for herself but also for P [the team composed by the two players]. But the limitations of her frame – her blinkeredness – preclude such considerations” (*ibid.*, 13; my emphasis). Accordingly a player in a basic frame cannot think

that her coplayers can reason in another way than her⁴⁵. To the contrary in S frames (“superordinate” frames) players conceive the situation both in “I/he” and “we” perspectives; they have “concept-clusters”⁴⁶ which encompass the two simple frames. They can handle the two “associated patterns of reasoning” (*ibid.*, 14). Bacharach claims that when an agent is in an “S frame” “she thinks about the possibility, and consequences, for her choice, of her coplayer’s in each of the subordinate frames” (*ibidem*)⁴⁷. Accordingly, a player in an “S frame” evaluates both group and individual payoffs (*ibid.*, 15)⁴⁸. In a way, reasoning within “S frames” allows players to evaluate the respective risk of reasoning individually or collectively, they can handle the “two cautious evaluations [,] *personal evaluations* and *group evaluations*” (*ibid.*, 16). Formally, Bacharach assumes a probability distribution over each frame in the set of players⁴⁹. Besides, Bacharach asserts that “a player in S [frames] has correct probabilities for her coplayer’s frame” (*ibid.*, 15). She knows the probability distribution and knows what the other players in each of these cases can play so her utility is traditionally a subjective (and individual) expected utility.

Nevertheless, for Bacharach this model does not fully explain why it is rational to team reason. Indeed, in “S frames” TR remains partly an individual and instrumental reasoning: “the present model fails in an important task, namely to show the possibility of rational we-reasoning. For it might be thought that when the agent is in the sophisticated frames, S, and looking at things from the group perspective, she is not we-reasoning, but simply guided as an individual by ‘group altruism’. She works out, on the hypothesis that her coplayer is in a certain frame and also will take a corresponding choice, what act of her own is best for the group interest. This sounds like ordinary instrumental I-reasoning with group regarding goals” (*ibid.*, 25).

The limit Bacharach identifies is explained by the fact that to fully adopt the TR, players must undergo an “agency transformation”. It is this “agency transformation” which leads to the TR (in Gold, Sugden, 2006: 90), and to the respecification of games⁵⁰. This “agency transformation” is the cornerstone of Sugden’s TR. The “unit of agency” in games is no longer individual players but the set of players as a team.

Quite obviously, players in “S frames” do not undergo this agency transformation. Besides, frames are supposed to be “non integrable”. Bacharach (1997: 25) emphasizes “the psychological impossibility of simultaneously seeing the problem from the ‘we’ and ‘I/he’ point of view”. This statement contradicts with the concept of “S frames” since they both entail “I” and “we” concepts in the players’ mind. Hence this concept is dismissed in the 1999 model.

In addition to the deletion of “S frames”, the 1999 model makes the relationship between players and the collective more complex. Multiple teams can compete and players can belong to different teams. However, each player is active in only one of the team to which she belongs. The underlying idea is the fact that with an important number of individuals interacting some of these individuals may fail to cooperate. It is an unreliable context. From the viewpoint of a player, because some other players may fail to be active in her team (as they belong and act for other teams), in order to maximize the collective utility of her team, she has to take into account who is indeed active; she has to consider “co-members’ unreliability” (Bacharach, 1999: 119). In this way, team reasoning becomes “circumspect” (*ibid.*, 118). A “signal” specifies for each player in which team she will be

⁴⁵ In fact, “the agent’s frame ‘blinker’ her: it is too narrow to enable her to see a highly relevant possibility. This possibility is that her coplayer is in the second simple frame” (*ibid.*, 14).

⁴⁶ Framing is portrayed like in the VFT: “a frame [...] is a cluster of concepts and associated pattern of reasoning, which an agent is aware of and uses in thinking about her situation” (*ibid.*, 3).

⁴⁷ This statement echoes Bacharach and Stahl’s late model “Variable-Frame Level-*n* Theory” (2000).

⁴⁸ Superordinate frames allow “modelling the choices of ‘cautious I/he-thinkers’ and ‘cautious we-thinkers’, not the unreflecting I/he-thinkers and we-thinkers of frames I and W” (*ibid.*, 16).

⁴⁹ “[A] player has frames W, I, S with probabilities ω , η , σ respectively, where $\omega+\eta+\sigma=1$ ” (*ibid.*, 15).

⁵⁰ “[H]aving the group interest at heart does not [...] suffice to explain [...] cooperation [...] something more happens [...]. This something is ‘agency transformation’. The key to my explanation is that agency transformation involves not only a transformation of payoff but also a transformation of reasoning” (*ibid.*, 90).

operative. More precisely, each player receives a signal telling her: (i) her “participation state” (i.e. the team in which she will be operative) and (ii) the probability distribution over the others’ “participation state”^{51,52}. The set of signals is given by an objective probability distribution (*ibid.*, 122). According to her signal, a player ““finds herself” participating in a certain team: her participation is not a choice, nor otherwise endogenized” (*ibidem*). And finally a function, labelled “a protocol”, ensures the link between players’ signal and their available and corresponding actions (*ibidem*). Basically, a protocol is a decision function⁵³. A team reasoner has to determine a “protocol” for her team, i.e. the set of individual actions (or the set of individual protocols) leading to the highest collective expected utility, and to do that, she needs to know what the other players will do, i.e. if they lapse to be active in her team, or not (if they lapse they will be active in another team) (*ibidem*). Consequently, an equilibrium in this model is a profile of protocols, which is optimal for the teams (*ibid.*, 124).

Following this idea of unreliability, Bacharach (in Gold, Sugden, 2006) identifies three types of contexts of team reasoning, with respect to the available information:

- i. In the “Simple coordination context”, the entire set of agents in the game adopts the team reasoning. If everyone in the set team reasons the outcome is Pareto optimal – the best profile in terms of the team utility function is reached (Bacharach, in Gold, Sugden, 2006: 122-123). A condition is required, all the players must know the payoff structure (*ibid.*, 123)⁵⁴.

The problem of failure leads to two other contexts inducing a “restricted team reasoning” and a “circumspect team reasoning”.

- ii. In the “restricted coordination context” several agents fail to team reason. In this case only a subset of agents team reason, the others are considered as the “remainder”. This time the team reasoners have to do “as best as they can between them, doing without the non-team reasoners” (*ibid.*, 127). Three conditions about the information structure are required: the players must know the structure of payoffs, the team reasoners must know the actions of the remainder, and they must know how many players constitute the remainder.

The epistemic requirements of this context explains why Bacharach asserts that “the restricted team reasoning still has two serious limitations: first, for it to be feasible the reasoners must know who is and who isn’t a team reasoner, while in practice there is very often great uncertainty about who is and is not governed by a choice mechanism” (*ibid.*, 130); and second, the choices of the remainder must be known (*ibid.*, 131).

- iii. The “unreliable coordination” context is a generalization of the previous one and is defined exactly as in the 1999 paper. There is a probability of failing and a determined profile of actions for those who failed: the “default choice” (*ibid.*, 131). In this context, team reasoning and more peculiarly the “circumspect team reasoning” is a valid way of reasoning only if all the members of the team know the payoff structure, the probability of failing and the default choice for those who failed (*ibid.*, 133).

⁵¹ Player’s signal gives them “the joint probability distribution of this signal and other agent’s states” (*ibid.*, 121).

⁵² Bacharach’s choice to refer to a “state” and not a “type” echoes his 1995 paper. He explicitly recognizes that “type” may be a confusing term in the following quotation: “[w]e can think of *i* as being one of two types in the sense of incomplete information game theory: participating, in which case she team reasons for *M* [the team], or lapsing. But since “type” may misleadingly suggest a permanent trait, I shall speak not of *i*’s “type” but for her “state” “ (*ibid.*, 121).

⁵³ Bacharach draws this concept on multi-agent systems (*ibidem*).

⁵⁴ The structure of payoffs gives all the required relevant information in order for players to team reason, i.e. the collective utility function, the associated collective profile of actions and in turn the individual components in the collective profile of actions. Normally Bacharach should have added a common knowledge condition, but he had not the time to do so (see Gold and Sugden’s footnote n. 5 [*ibid.*, 151]).

3.2. Is cooperation naturally or “interactionally” based? How can multiple selves be conciliated?

A tradeoff between standard reasoning and team reasoning exists in every models of TR proposed by Bacharach. I suggest that this struggle stems from two intertwined aspects underlined in his theory of TR. First, players possess a predisposition in a way or another to cooperate, like if it was in their ‘background’. This implies that even if the games’ outcome relies on strategic interactions, cooperation is weakly explained by these interactions. In 1995 cooperation relies on natural types. In the other models it is explained by cognitive and psychological *predispositions*⁵⁵. Some of the time, i.e. depending on the game, collective cognitive representations take hold over individualistic ones. These presuppositions imply that Bacharach has to justify why players (or some of the players) who are *predetermined* to cooperate, do not. Second, players can have dual or multiple selves (depending on the models) and Bacharach has some problems to reconcile players’ competing selves.

This tradeoff between (i) the two modes of reasoning, and (ii) players’ competing selves, raises some difficulties and, again, there are different methodologies to bypass them⁵⁶.

In the 1995 model, two facts compete. On one hand the use of the term “natural types” suggests that players’ capacity to be cooperative is an intrinsic capacity, a quasi-genetic characteristic. Players of type T are, we could say, predetermined to cooperate. On the other hand, players do not cooperate invariably. To justify this, Bacharach (1995: 17) confesses: “the type of reasoning I have been modeling may become detached or abstracted from the specific natural types which were its seeding ground”. Two consequences follow: (i) Bacharach requires to an “outward sign”, minimizing the capacity of types T to recognize the other types T, to justify that some of the time types T do not cooperate with other types T⁵⁷; and (ii) when types T encounter best-reply reasoners, they must play *against* their natural tendency, since they are *forced* to adopt a standard individualistic reasoning, their behavior going against their nature. It is in such ways that players are “abstracted” from their natural types. For Bacharach the problem of this model can be resolved by the integration of framing: “[a]n alternative development would make T membership a variable element in players’ frames in the sense of variable frame theory (Bacharach, in Binmore et al., 1993): that is a player might or might not think about the game in terms of whether she and her coplayers belong to T. In the case in which T is the player set, we may put this by saying that a player may or may not think in ‘we’ terms about how to play the game. The more inclined a player is to ‘we’ thinking, and the more inclined she takes coplayers to be, the more will fellow-member reasoning be favoured” (*ibidem*). Furthermore, I would like to shed lights on the fact that postulating that a player cooperate this or that person *a priori* because of mutual recognition between the players who possess the same natural type does not fit with a game theoretical framework.

To bypass the problem which Bacharach identifies, the purpose of the 1997 model is to justify *via* a “variable frame model” when players cooperate and when they fail to cooperate, i.e. when they team reason or remain individualistic reasoners. In this model, it is the structure of games which enhances the frames associated with the two modes of reasoning. Hence the game may enhance pre-existing collective cognitive representations. If players conceive themselves as members of a team they are naturally led to cooperate, it is again a kind of genetic or naturalistic programme leading them. Nevertheless, another problem appears. Games can induce a tradeoff between cooperation

⁵⁵ We must refer to Scazzieri’s discussion of framing in Bacharach’s VFT (cf. 2.2). In fact, what is true for the VFT is true for TR, since the mechanism of framing identified by Bacharach both in VFT and TR is the same.

⁵⁶ Tan and Zizzo (2008) grounded on Bacharach’s TR propose a way to escape from these pitfalls. Again, Sugden (1993, 2000, 2005, 2006) is not concerned by this, as his TR relies on another approach, more conform to decision theory.

⁵⁷ According to Bacharach “this sign is not a signal: T’s do not decide whether or not to display it” (*ibid.*, 11). This idea echoes the 1999 model.

and defection. In fact, even if coordination games⁵⁸ show scope for cooperation (or coordination) on the Pareto best outcome, the riskiness of cooperation may induce individualistic behavior. In this perspective, the “agency transformation” allowing players to team reason and then to cooperate, may not be enhanced. Moreover, this tradeoff is internal for players in S frames when evaluation of cooperation in a I or a we perspective do not lead to the same conclusions.

Subsequently, Bacharach refines the conditions enhancing we frames in the later models (1999, 2006), partly to avoid this pitfall (the other reason will be exposed in the next subsection). Based on the literature on social psychology and to experiments within this subfield of psychology (Bacharach, in Gold, Sugden, 2006: 132), Bacharach precisely defines the conditions and the *mechanism* which trigger we frames⁵⁹. For instance, experiments in the theories of self-identity and self-categorization, in social psychology, point out that “common fate” or “common interest” and “interdependence” (among others), induce group-identification. Bacharach (*ibid.*, 82) argues that “‘interdependence’ and ‘common interests’, are of particular interest for our inquiry because they connect group identity to characteristics of decisions problems”. They “may help explain why a group of people locked into a game might tend to group identify. They have a common interest in Pareto-optimal outcome and, usually, they had no hand to in choosing the payoff matrix. I shall call the tendency of this combination of factors to produce group identification in the player set the common problem mechanism” (*ibid.*, 133). In particular, the game theoretical characteristics at sight in this “common problem mechanism” are the following:

- i. “ P_1 [Player 1] and P_2 [Player 2] have a common interest in s^* over s [s^* and s being two different outcomes] if they mutually know that both prefer s^* to s ” (*ibid.*, 83),
- ii. “among all the feasible outcomes there is one – call it s^* – both P_1 and P_2 rank highest, and that it can only be brought about by each acting in a particular way. This is the pattern in a ‘common-interest’ game” (*ibid.*, 84),
- iii. and finally “the interdependent hypothesis concerns situations in which s^* is not assured by individualistic decision-making, and P_1 and P_2 perceive that it is not” (*ibid.*, 85).

Accordingly, group identification is produced if “[f]or some S, S^* the players have common interest in, and copower for, S^* over S , and $sol(G)$ [i.e. the set of solutions in the game G] contains outcomes in S^* ” (*ibidem*). However, in addition, players must have common knowledge of these facts, i.e. of the “common interest”, “copower”, and “interdependence” characteristics (*ibid.*, 83).

Although, once more, players may not perceive these characteristics, they may not perceive this “common problem mechanism”. Thus, “we frames” are not necessarily salient and subsequently enhanced.

All of these difficulties in the justifications of cooperation must be considered alongside with the competing selves of players. In 1995, Bacharach hardly reconciles the fact that “types T”, which are naturally cooperative reasoners, may act as “BR-reasoners”. As stressed above, Bacharach needs to require to an *ad hoc* assumption contradictory to his conceptual choices. In the subsequent models of TR, as soon as TR relies on framing, the incompatibility between players’ two selves is even more exhibited. Indeed, Bacharach highlights and advocates that frames are “non-integrable”. If they are “non-integrable” it means that there is no link between players’ different selves, either they are a member of a group, or they are an individual. In the presentation of the 1997 model Bacharach (1997: 2) emphasizes: “[t]he ‘we’ and ‘I’ ways of thinking about a problem are rivals”. The concept of “S frames” is, to some extent, an attempt to reconcile these two competing selves. Nevertheless, players in S frames cooperate in their own interest and we see an incompatibility

⁵⁸ Coordination games as defined by Bacharach in this paper correspond to a broad category which encompasses Prisoner’s Dilemma, Stag Hunt (both Rousseau’s Stag Hunt and Aumann’s Stag Hunt), the Assurance Game as defined by Sen, and finally the Hi-Lo Game.

⁵⁹ Bacharach is well aware of the literature on social psychology and the debates between psychologists and especially between cognitivists and the “interdependence theorists” who do not agree on the different conditions which prone a sense of group-identity in players’ minds. For more details, we must refer to the pages 77 and 78.

between cooperation as profile directed reasoning induced by a “collective profile function” and evaluating risk to cooperate from an individual perspective. The tradeoff between team reasoning and traditional reasoning when a player is in a “S frame” is resolved by a principle of “reason dominance”⁶⁰, explained by a principle of “payoff dominance”. Bacharach explicitly recognizes that this concept is a limitation for his model: “[i]n the absence of a natural way to resolve conflicts, *we are limited once again*, in making predictions from our theory, to use of dominance criteria” (Bacharach, 1997: 17; my emphasis). Nevertheless, in this conception, an even more embarrassing problem is that, to evaluate what is of interest for them players in “S frames” must have an individual utility function. They do not follow the team profile of action. They must evaluate themselves the action to take, so they refer to a proper subjective utility function. Hence even if they may choose to cooperate – as they can in “we frames” – they remain individualistic reasoners⁶¹. So again, we see the incompatibility between collective and individual selves. Bacharach is aware of this fact, as he recognizes that he fails to demonstrate that cooperation, i.e. we reasoning, is rational. This methodological ambiguity is dismissed in 1999 and then in 2006. In addition the 1999 model offers an original account of players, even within Bacharach’s work. Players can have multiple selves with respect to the fact that they can belong to several teams (Davis, 2011: 118)⁶². Nevertheless, once again, only one self can be enhanced at the same time, since players can be active in only one of the teams to which they belong. When a player has to take a decision and when this decision is acting for one of the team to which she belongs, she fully endorses the collective self, inherited from the team. Thus, again, there is the same idea of “non-integrability”, this time, between multiple selves and no longer between dual selves. Players do not have the reflexive ability to choose between competing aspects of their identity.

Another question tightly linked to the competing selves, is how far can Bacharach go in the tradeoff between an individual’s interest and the collective interest? In other words, how far can individuals go in their renouncement of their individual identity when they belong to a team and act for that team?⁶³ Except in his 1997 paper the question of the competition between individual and collective interests is never explicitly posed. Bacharach argues that individual and collective interest may not be concomitant⁶⁴. This postulation is clear when he asserts that “the goal of the group in this framing of self need not agree with the person’s goals under her individual framing of herself, but perceived agreement of individual goals among a set of individuals favours framing as members of a group with this common goal” (Bacharach, in Gold, Sugden, 2006: 90). Though, the competition between “I frames” and “we frames”, or between individualistic reasoning and team reasoning, is obviously the proof that if individual and collective interests compete, team reasoning may not be enhanced. In addition, when players belong to and act for a team, Bacharach implicitly presupposes a kind of convergence within the members’ representations and in turn in their actions. In fact, Bacharach is much more insistent in this book about the “common ranking” condition. Team reasoning is induced by condition in which this common ranking condition prevails: “the basic class

⁶⁰ Bacharach defines dominance as follows: “z *strongly reason-dominates* z’ if it is better than z’ in term of both group evaluation and personal evaluation in S. If z strongly reason-dominates z’ [z and z’ are two possible actions, e.g. cooperate or defect respectively], then in cases in which personal and group evaluation are comparable, z defeats z’ on balance *whatever* the relative weights. z *weakly reason-dominates* z’ if it is better than z’ in terms of some evaluation, and not worse than z’ in either” (Bacharach, 1997: 17).

⁶¹ For Bacharach, players in “S frames” have basically altruist preferences (*ibid.*, 25).

⁶² It is interesting to note that in this account, players can be heterogeneous. As soon as they belong to different teams they potentially have different goals. In this perspective, it is different from standard non-cooperative game theory in which everybody is structurally equivalent (i.e. every player has the same set of strategies and the same utility function). Notwithstanding, inside each team, every member has the same purpose: maximizing their team objective function, which is the Paretian ranking of their aggregated individual utilities.

⁶³ This questioning echoes Gilbert’s account of collective entity (e.g. 2003). For Gilbert, when individuals accept to enter in a collective action, they become entirely committed to the collective purpose and then become collective entity. They cease to be fully-fledged individuals.

⁶⁴ In fact, Sugden shares this point of view, we can refer to his example in 1993.

of situations in which the possibility of team reasoning arises is that there is a set of agents who have alternative action options, and a common ranking by the agents of the profiles made up by these alternatives. I call situations of this kind coordination contexts” (*ibid.*, 122). Subsequently, on one hand, the players’ individual interest and the collective interest must be concordant, and on the other hand, players’ interest must be convergent to enhance TR. Besides, Bacharach’s models rely on a *principle of aggregation*. The collective utility function is simply the addition of the individuals’ utilities and Bacharach admits that “there are well known difficulties in aggregating values in general” (Bacharach, 1997: 17). Bacharach is aware that this mechanism is quite controversial to grasp the problem of what is a collective. In fact, it implies that players’ individual interests are perfectly concordant to the collective ones. Practically, Sugden faces the same problem the collective utility function is the sum of the individuals’ utility functions. In reality this trouble is internal to a standard game theoretical framework. It is neither specific to Bacharach, nor to Sugden.

In summary, throughout Bacharach’s work the explanation of TR and accordingly, cooperation, seem to rely on a natural propensity possessed by players who are akin to cooperate with individuals who belong to the same group as them. The importance of the environment is not dismissed, since framing relies on the environment, and shall we say, on the interactional context. Nevertheless, it is not through interactions that, at certain points in time, players decide to cooperate or to act collectively as a team. They do not agree on a mode of interaction. They possess *a priori* some propensities to act in one way or another. It is given to them; and sometimes, as a theorist Bacharach decides that the game’s conditions can conform the requirement to trigger cooperative behaviors. This statement contradicts Bacharach’s attempt to endogenize the probability that a player, in a given context, perceives herself as a member of a team and no longer as an individual. Moreover, the endogenization problem exemplifies the tradeoff between the individual self and the collective self of players.

3.3. Salience and the “endogenization problem”

As stated above, the incompatibility between the players’ different “agency” must be considered alongside with the salience of frames and the endogenization problem. If collective identity is more salient than individual identity, team reasoning and consequently cooperation takes hold over individualistic reasoning and potential defection. This problem of endogenization arises with the inclusion of frames in team reasoning. In fact, this difficulty is not entailed in the 1995 model. However, as soon as frames matter, Bacharach’s purpose is to endogenize the probability that in a given context “we frames” get the upper hand on “I frames”. This endogenization problem explains the progressive refinements of the conditions that trigger “we frames”, and the issuing concept of “common problem mechanism”.

As Bacharach claims (in Gold, Sugden, 2006: 81): “I am interested in the effects of spontaneous framing in games. Group identification is a framing phenomenon. So I am interested in the role of spontaneous group identification in decision-making”. Besides, he affirms: “which of my collective personae is activated depends on the current ‘accessibility’ of the categories to which I belong [...] In human interactions, the accessibility of categories is a special case of the notion of availability of frames at the heart of the variable frame theory of games. The process in which categories are activated is context-dependent and jumpy” (*ibid.*, 74). Consequently, the problem of endogenization and salience are tightly linked. To avoid the problem of endogenization, Bacharach needs to explain how and when collective identities may be more salient, i.e. in which kinds of games or interactional contexts.

In 1997, Bacharach proposes two models: (i) one without players in “S frames” and (ii) one with players “S frames”. In the former if players are in “we frames”, it is *a priori* postulated, and not

triggered by the game. Besides, players in “we frames” are “blinkered” by their frames. They cannot think of the others’ unreliability induced by the relative risk of cooperation. This relative risk does not compete with the enhancement of “we frames”. Players are already cognitively predetermined to cooperate. Accordingly framing is not a matter of games and salience. In the latter, the respective probabilities of being in I, W or S frames are given by an objective probability distribution over the set of players, whereas Bacharach (1997: 6) argues that “collective identity effects are [...] endogenous to [games]”. In fact, “the personal and social levels of categorization have varying different salience in different interactions” (*ibid.*, 23); these probabilities should be determined by the games’ characteristics, and particularly by what Bacharach calls “the ‘gain from cooperation’” (*ibid.*, 24). The probability that “we frames” come to the players’ mind may increase with this “gain from cooperation” or may be related to what Bacharach calls “the harmony of interest in game” (*ibidem*). If games set out these two dimensions, i.e. “gain for cooperation” and “harmony of interest”, they are games of “common interest” (*ibidem*).

The same arguments about the games’ characteristics tending to promote we frames are developed in 1999, and once again Bacharach emphasizes his failure. He identifies some paths to resolve the problem: “it is plausible that ω [the probability that “we frames” instantiate] may be an increasing function of certain quantitative features of the payoff structure such as “scope for co-operation” and “harmony of interest” [...]. To endogenize ω [...] one must show that the payoffs and other constitutive features of the basic game make collective identity salient or otherwise tend to induce team-thinking” (Bacharach, 1999: 144)⁶⁵. So here again Bacharach is not able to bypass the limits he encounters in the previous paper; even by changing some methodological and formal aspects of his model.

Even with a precise identification of these characteristics, in his posthumous book, the problem of endogenization still prevails. Indeed, due to the incompleteness of players’ frames, players may not perceive salience of the “common problem mechanism”. Group-identification as a result may not be prompted (in Gold, Sugden, 2006: 87). Bacharach explains that the characteristics underlying the “common problem mechanism” are more noticeable in some games than in others, as is group-identification (*ibidem*). Bacharach gives the following example:

“In prisoners dilemma, players might see only, or most powerfully, the feature of common interest and reciprocal dependence which lie in the payoffs of the main diagonal. But they might see the problem in other ways. For example, someone might be struck by the thought that her coplayer is in a position to double cross her by playing D [i.e. Defection] in the expectation that she will play C [i.e. Cooperation]. This perception might inhibit group identification” (*ibid.*, 86).

In fact, Bacharach argues:

“There is common interest in Stag Hunts, Battles of the Sexes, bargaining games and even Prisoner’s Dilemmas. Indeed, in any interaction modelable as a ‘mixed motive’ game there is an element of common interest. Moreover, in most of the landmark cases, including the prisoner’s dilemma, the common interest is of the kind that creates strong interdependence, and so [...] creates pressure for group identification. And given group identification, we should expect team reasoning. But for the theory of endogenous team reasoning there are two differences between the Hi-Lo case and these other cases of strong interdependence. First outside Hi-Los there are counterpressures towards individual self-identification and so I-framing of the problem” (*ibid.*, 144).

The second difference is about the “unanimity condition”. Saying that there is a unanimity

⁶⁵ Tan and Zizzo (2008) build their model around the concept of “harmony of interests” and “disharmony of interests”. They dissociate games by showing that some games like Hi-Lo games show an “harmony of interest” and others, like the prisoners dilemma show the opposite, i.e. a “disharmony of interests”. However, contrary to Bacharach, they succeed in endogenizing the probability that a given player in a given game fails to team reason.

condition implies that players “have the same value on every profile in the profile set” (*ibid.*, 88).

I would like to point out that the three types of contexts identified in 2006 (in Gold, Sugden) could induce an ambiguity on what can affect salience of “we frames”. If we interpret the three different coordination contexts as an informational device (potentially due to the different size of the set of players), it could challenge the interpretation of salience with respect to the structure of objective game matrices. Consequently, two facts could potentially affect the enhancement of the players’ collective identity: (i) the types of games, and (ii) the risk of cooperation according to the epistemic structure of games. Subsequently, it seems that the context of coordination is a question of information and knowledge and not only of unreliability *per se*. So in the end we do not know if this is due primarily to the type of games, which entails this or that type of interactions and the underlined unreliability, and accordingly the different saliencies of frames, or to the size of the set of players and the lack of information. Nevertheless, as the distribution of cooperative and non-cooperative players is already done before the game and symbolized by an objective probability distribution, a problem remains to justify that it is the uncertainty, regarding the available information, which affects salience.

Finally, I would like to highlight that the problem of endogenization, contrary to aggregation, is not specific to game theory but to Bacharach himself and his methodological choices. If the aggregative problem must be considered alongside to standard game theory and neoclassical economics (in their approach of collective entities and collective states); the former difficulty, explaining Bacharach’s progressive refinements in his theory of TR, is totally specific to the integration of framing in a game theoretical account. His failure can be explained by the incompatibility of Bacharach’s attempt with standard non-cooperative game theory. In fact, as the structure of the game must be known prior to the game in order for players to have beliefs about the others’ type and corresponding behavior, there is an impossibility to endogenize the salience of frames. Moreover, as previously emphasized, Sugden does not encounter this pitfall. He grounds his theory of TR on the same fundamentals of decision theory, thus this problem is inexistent for him.

I suggest that, like in VFT, Bacharach faces numerous complications because he assumes two positions at the same time, that of the players and that of the theorist. He tries at the same time to build his theory of TR putting himself in the position of a player, but establishes initial requirements as a theorist, in order for his models to be defined. We obviously do not face the same dual position in Sugden’s account of TR. As a theorist he postulates the requirements of TR to be a valid reasoning, and the requirements of his theory to be “internally consistent” (Sugden, in Gold, 2005: 196).

Even if I have insisted on Bacharach’s problems induced by his conception of TR and the subsequent conceptual and methodological variations enhanced, I would like to highlight in the next section that it does not mean that his program fails, quite the contrary. I shall emphasize how, through both VFT and TR, he changes the standard conception of players and their rationality. In addition, these enrichments of the standard conception of players and their rationality are rarely proposed in such a way. Nevertheless he has had to pay a very high price, if we compare with Sugden’s TR which shows how it is possible to enrich standard game theory without endangering the model proposed. Conversely to Bacharach, Sugden’s TR is perfectly logical and consistent and is a *quasi-closed system*. The only difficulty enhanced in Sugden’s TR is the justification of “the assurance condition”, i.e. the “common reason to believe” that each player or team member “endorses” TR. Indeed, this “condition of assurance” relies on something which stems from outside the games. Sugden (in Gold, 2005: 196) acknowledges: “[t]hat assurance might be induced by players’ common experience that, in some analogous class of recurrent interactions, people like them do in fact engage in team reasoning. But, in the absence of common experience of team reasoning, and in the absence of any prior agreement to engage in it, assurance may not be

possible”.

4. A rational reconstruction of VFT and TR’s enrichments of standard non-cooperative game theory: a new conception of players and their rationality

In this section I shall highlight how Bacharach with VFT and TR progressively enriches the standard conception of players and their rationality. In this manner VFT and TR must be considered as a continuation enhancing different but progressive enrichments participating in a same purpose. Bacharach has always purported to justify why *it is rational for players to coordinate and cooperate*. He attempts to offer a game theoretical explanation of empirical regularities demonstrating that players are more inclined to coordinate and cooperate than to defect. If VFT is a means to justify coordination on salient options, TR is a means to justify cooperation. Nevertheless, the objective remains the same, i.e. to offer *a more realistic account of interdependent decisions in strategic contexts*.

I shall show how this purpose translated in his work entails a different conception of players compared to standard game theory and then a different account of rationality in games. Indeed, Bacharach’s attempt was first and foremost to formalize more realistic players (Gold and Sugden, 2006), and more realistic players issue, as I shall try to show, in a social and collective account of players’ rationality. Asserting that players’ rationality is social means that the determinants of players’ choice are not entirely explained by internal consistency of choices but by the context, by cultural determinants, and partly by the others with whom they are interacting, their culture, etc. For instance, players’ preferences regarding games’ solutions depend on others’ representations or frames. Now, asserting that rationality is collective means that players have a collective objective necessitating interdependence and knowledge of this fact. Players no longer act for themselves as individuals but for the group. In Sugden’ words (in Gold, 2005: 183): “[t]he rationality of each individual’s action derives from the rationality of the joint action of the team”.

4.1. Which conception of players?

Bacharach regularly refers to ‘normal’ agents (1991, 1993, 1997). I shall emphasize that this conception of players is mostly based on three dimensions. Even with various models of VFT and TR, different conceptual and methodological choices, and different formalizations, we always find three concerns: (i) he tries to avoid to endow players with unlimited reasoning depth, (ii) he integrates psychological dimensions in the way players frame their decision problem, and (iii) he tries to justify how players’ identity matters in decision problems, even for strategic decisions. Obviously they are tightly linked in Bacharach’s works and sometimes it can be hard to figure out which one has been improved.

First of all, Bacharach’s conception of ‘normal’ players entails limited cognitive abilities. This hypothesis relies on two distinct but related aspects: the depth of players’ beliefs and the epistemic requirements of *rational choices* and then, conceptual limitations. As Bacharach (1992: 247) already claims, “there are limits about the depth of the beliefs people can attain”. Consequently game theorists should not explain games solutions using common aligned beliefs and common knowledge. For instance, to the question “[w]hat difference does common knowledge make?” he answers “in many cases, none; ck [common knowledge] itself is not crucial – all that matters is whether there is iterated knowledge of sufficient degree” (Bacharach, 1992: 10). Accordingly, in his skeletal development of both VFT and TR (in 1991 and 1995 respectively), one

of his major concerns is the strength of players' beliefs. They are both grounded on a "sufficiently strong degree of belief" to avoid the unrealistic common knowledge requirement (Bacharach, 1991: 30; 1995: 9), and again it is due to players limited cognitive abilities. In summary, Bacharach generally avoids adhering to the common knowledge hypothesis throughout his models of VFT and TR. Even in 2006 (in Gold, Sugden), when Bacharach states his "common problem mechanism" on common knowledge, he tries to assess team reasoning in contexts characterized by a lack of information (we must refer to the three types of coordination identified).

Another aspect of players' limited cognitive capacities is explained by their psychology, and more precisely by their perceptions. In his early works Bacharach does not mention psychological assumptions about players' reasoning so he only refers to knowledge, but as soon as he begins to work on framing, players' cognitive limits are first and foremost due to *incompleteness of their perceptions*. Indeed, Bacharach argues, "our conceptual endowments are limited, that we as normals are not 'conceptually omniscient'" (1991: 29). It means and implies that players' *frames are incomplete*. For instance, Bacharach asserts that there are *universal frames* encompassing all the possible representations of a given problem but players, as 'normal' individuals, cannot have access to these universal frames. They cannot perceive all of the relevant ways to represent and describe a decision problem (1991, 1993, 2001, 2006). Thus, as Bacharach refers to incompleteness of framing – or in other words to incompleteness of representations, by direct link it means that the players' information is incomplete. Besides, since framing is anterior to reasoning, there are necessarily limits about the depth of players' reasoning⁶⁶. Players' awareness of all their own possible available strategies and, by implication, of others' strategies are necessarily incomplete. Hence players' beliefs are necessarily limited and incomplete. Moreover, in most of his work knowledge is an "occurrent knowledge", i.e. conditioned by players' frames; and because frames are necessarily incomplete, so is players' knowledge. Bacharach (e.g. in 1991, 1993) regularly insists on the fact that, due to frames, the incompleteness of players' information (or knowledge) is deeper than Harsanyi's (1967/8). For instance Bacharach (in Binmore et al., 1993: 257) claims: "[q]uite generally, a given question may or may not occur to an agent facing a decision problem; she may or may not think of it. This phenomenon gives rise to variations in agent's beliefs systems of a more radical sort than those to be found in the familiar Bayesian model". In addition, Bacharach presupposes that culture induces biases of representations. He refers to an empirical example: the Eskimos have numerous ways to describe the snow and the color white which we are not aware of as Europeans or more generally as non-Eskimos (*ibid.*, 259). In this perspective, individuals and subsequently players, are in a way culturally determined and accordingly cognitively determined. This cultural determinism may induce again limited cognitive capacities. It is precisely both the limited representations and this determination which allow players to be successful in VFT.

I would like to highlight that Bacharach's integration of framing in his theories portrays players quite differently from framing in the literature on this topic, and especially from Kahneman and Tversky (1979, 1986) to whom Bacharach refers. Framing does not entail irrational players for Bacharach as is generally assumed. It is quite the contrary frames allow Bacharach to nicely shape what could be understood as an idiosyncratic individual rationality. By arguing that deliberation, i.e. the phase during which rationality is operative is distinct from the unconscious phase, he shows that framing does not entail irrationality.

Furthermore, frames give access to a part of players' identity if we refer to Scazzieri (in Gavalotti et al., 1997). For instance, he emphasizes that in Bacharach's account: "framing would primarily be

⁶⁶ "[W]e can't think about everything. One reason is that, because every real situation has infinitely many aspects and/or we are bounded, there are always aspects of a situation of which we are unaware. That is, always frames are incomplete [...]. One consequence of this, and the logical priority of framing to believing, is that the space of propositions or event on which an agent's subjective probabilities are defined is always incomplete" (Bacharach, 2001: 5).

associated with the cognitive and linguistic ability to grasp specific problem situations through the activation of a particular set of ‘naturally connected’ features. It is reasonable to assume that frames would often be associated with relations among attributes, and that such relations will often be of the causal type. In short, framing would generally rely upon pre-existing cognitive structures (the different subsets of naturally connected features), but only specific (contingent) circumstances could turn a virtual frame into an effective one” (Scazzieri, in Gavalotti et al., 1997: 196-97). Accordingly frames rely on both the players’ culture and personal experience. In Bacharach’s theories, players are not discarded from what comprises their specificity, contrary to standard non-cooperative game theory.

TR, relying on framing too, sheds lights on another aspect of players’ identity. When players team reason, i.e. when they are in we frames, they express their “group identity”; they “[think] of [their] properties as the subset of the properties which are characteristics of members of that group” (Bacharach in Gold and Sugden, 2006: 73). From the viewpoint of a player, seeing herself as a member of a group is a process of “affiliation” which is “a psychological process in which a person who does think about a certain group, defined by some shared property, comes to think about it as ‘us’” (Bacharach, 1997: 2). When a player “identifies herself with the group; her self-conception is as a component part of the group. This is reflected in her language. She thinks and speaks of the group not as “them and me” but as “us”” (Bacharach, 1999: 134). Being a member of a group and acting for this group implies in her mind a “cognitive extension [...] of self-interest” meaning that “the group’s goals define her basic preferences” (Bacharach, 1997: 16). Accordingly we frames and TR express players’ “collective self-conception”. Bacharach emphasizes that a “personhood is to some extent constituted by group membership [...] personhood is the resultant, to the extent that it is so constituted, of a set of group identities; more exactly, the person is defined by the intersection of her group identities” (Bacharach in Gold and Sugden, 2006: 88-9). As a consequence “team reasoning is a basic decision-making proclivity and mankind” (*ibid.*, 121). Moreover, Bacharach states his conception of the connection between collective and individual agency on psychologists within self-categorization theory and more specifically on Tajfel and Turner (1985) who assert: “the sense of group identity precedes, developmentally, the sense of personal identity”.

In the previous section I underlined that Bacharach’s models of TR depict players as dual and multiple selves. I insisted on the recurrent tradeoff between individualistic reasoning and team reasoning, and the incompatibility between them. The difficulty I emphasized was primarily theoretic. I do not mean that Bacharach’s conception of players as individuals may not be consistent, or that Bacharach fails to demonstrate what a player is. I merely suggested that such a conception of a player raises numerous theoretic difficulties and that it largely endangers the validity of Bacharach’s models. In fact, I agree with Davis’ (2011: 118) point of view when he asserts that in Bacharach’s TR “having multiple selves underlies being a single individual”. This assumption is reinforced even more by the previous quotations. According to him, “Bacharach sees individuals’ multiple selves as their supra-personal social identities. Thus to show that people are single individuals, he needs to show that they do not dissolve into the many different social groups with which they identify” (*ibid.*, 119). Bacharach’s challenge is accordingly “the fusions of agency” (*ibid.*, 119). Bacharach’s modus operandi is the following for Davis, he “employ[s] one kind of relational conception of the individual – one in which people are single individuals in virtue of how their interaction with others makes them more than collection of multiple selves” (*ibid.*, 119). This is true if we consider the fact that only one mode of reasoning can be enhanced at one time.

4.2. “Game-theoretic rationality”: challenging the *individualism* postulate

In some relevant aspects both VFT and TR can be considered as “revisionary theories”, they are not “bounded rationality theories” (Gold Sugden, 2006; Sugden, 2001). For instance, “[i]n VFT, though players are *conceptually bounded* by their frames, they are not bounded in their strategic thinking within those frames” (Bacharach, Bernasconi, 1997: 39). To the contrary, both VFT and TR offer a revision of individual rationality in strategic contexts. Players’ choices are no longer contained in an individualist account. Both VFT and TR’s conceptual and methodological characteristics and the profound modifications of the standard conception of players they induce, have necessarily consequences in terms of rationality. However, some aspects of rationality in both of them remain quite familiar within a game theoretical framework. It is my opinion that a certain semantic of standard individual rationality prevails whereas its content sensibly changes in these theories. In fact, I would like to draw attention to the fact that by using VFT and TR Bacharach challenges the confusion between two distinct aspects of individual rationality which became confounded and assimilated in decision theory: (i) consistency of choices and (ii) maximization. In the VFT players explicitly maximize their subjective expected utility. Besides, “Bacharach interprets the payoffs of a game as specifying what the players want to achieve as individuals (or what counts as success for them as individuals). He assumes that payoffs can be treated as utility indices in the sense of expected utility theory so that, in situations of uncertainty, a player’s success is measured by the expected value her payoff” (Gold, Sugden, in Gold, Sugden, 2006: 8)⁶⁷. However it is the content of their utility and, shall I say, of their preferences, which is not standard. The objective of a team and the team members is again, to maximize the expected collective payoffs, i.e. to maximize the expected collective utility.

What matters for Bacharach, in game theory, is to formalize “practical” modes of reasoning, i.e. “what it is that makes actions “satisfactory” and how “satisfactoriness” relates to choice” (Bacharach, 1987: 17). Besides, a mode of reasoning must be “valid” for Bacharach: “we might call a mode of reasoning in games valid if it is ‘success-promoting’: given any game of some very broad class, yields only choices which tend to produce success, as measured by game payoffs” (Gold, Sugden, in Gold, Sugden, 2006: 8). This principle translates both in VFT and TR. In VFT, Bacharach postulates principles of rational choices (cf. 2.1) that any player must follow to be rational. Players must respect these principles in any strategic context, i.e. in any type of games. Regardless of the subjective representations they hold, successful reasoning, i.e. rational reasoning (according to Bacharach), relies on the principles of rational choices and the equilibrium selection which Bacharach edicts. This is the only way to reach the highest payoffs. In the TR, if players team reason they endorse the objective of their belonging team, i.e. to reach the highest collective payoff. Collective rationality is accordingly “success-promoting” too. From the same perspective as VFT, team reasoners have a determinate mode of reasoning, regardless of the coordination context (respectively to the lack of information) or the type of games (i.e. in prisoners dilemma, or Hi-Lo game, for instance). Therefore, Bacharach remains anchored in what Sugden (1991) calls the “principle of rational determinacy” (in a weakened form), which means that there is only one way to play rationally in all types of games. Both in VFT and TR, if players face the same information, they draw the same conclusions, using the same rules of inference.

Nevertheless, for numerous reasons standard individual rationality and the corresponding postulate of *individualism* are challenged. In order to understand the changing aspects of players’ rationality both in VFT and TR, I shall first present Bacharach’s early thoughts about the standard conception of rationality in game theory. He identifies, with Hurley, in the introduction of the book *Foundations of Decision Theory* in 1991, what he considers to be the *misspecifications* of individual decision theory. By doing so, they propose few ways to *revise* this standard account of rationality. I

⁶⁷ See the footnote n. 8 (in Gold Sugden, 2006: 32-33), in which Gold and Sugden explain and justify this postulate.

shall highlight how these early thoughts translate in his VFT and TR. Bacharach and Hurley investigate five dimensions of decision making, of which three are particularly relevant to understand the *semantic* and the *content* of rationality in VFT and TR: (i) “the structure of attitudes”, (ii) “humanity” and (iii) “individualism”.

The first concept “includes questions about the restrictions on the attitudes of the decision-maker – her beliefs and preferences, subjective probabilities and utilities – that rationality imposes” (*ibid.*, 1). They argue that the problem is that “[p]references are often sensitive not just to the intrinsic characteristics of the good but also, [...] to the context in which the intrinsically characterized good is located” (Bacharach, Hurley, 1991: 12).

In VFT players’ preferences are not “context-independent”, since they rely on players’ subjective representation which are context-dependent. As a consequence, the axiom of extensionality⁶⁸ is violated, and players’ choices do not respect the axiom of “consistency of choices”⁶⁹. For the rational choice theory it means that players are irrational. Naturally, this is not Bacharach’s point of view. In addition players’ preferences depend on their beliefs about others’ subjective representation. The options a player ranks higher than others depend on her beliefs about others’ perceptions. In standard game theory preferences are totally distinct from players’ beliefs about others’ choices or beliefs (Hargreaves Heap and Varoufakis (2004 [1995])). In the TR, players do not follow their individual preferences if they team reason. It is the collective profile of actions of the team which is reason giving. As emphasized above, the rationality of players derives from the rationality of the team (which again means reaching the highest collective payoffs).

The “humanity” of a decision-maker encompasses “her boundedness and her culture” (*ibid.*, 2). Bacharach and Hurley assert that, even if we suppose an idealized rational player (endowed with unlimited cognitive capacities and perfect knowledge), it may not be sufficient to lead to the determination of solutions in games: “even an extreme idealization of inferential powers fails to deliver the goods in the shape of determinate solutions to games” (*ibid.*, 2-3). As a consequence, “if we are to avoid the conclusion that rational action in games is impossible, it may be necessary to move away from the conception of rationality of traditional decision theory toward a more naturalistic conception in which “ought” implies “can”” (*ibid.*, 3). Taking into account the players’ humanity is one of the possible solutions. Indeed, the two characteristics at stakes in players’ humanity translate in both VFT and TR. Both in VFT and TR, players are conceptually bounded (their frames are necessarily incomplete). Furthermore and as developed above, players are culturally determined. In VFT, their representations are influenced by cognitive predispositions inherited from their experience and primarily by their culture. In VFT, it is precisely because players are cognitively bounded that they may be successful to resolve a game (1991, 1993, 1997).

Finally, Bacharach and Hurley, clearly stress that the “individualism” postulate, mainly in game theory, is question begging. They claim that “a number of questions arise about the relationship between individual rationality and game-theoretic rationality. We usually think of individual decision theory as in a sense prior to game theory, but as several points [...] we are led to question whether games may be embedded within supposedly individual decision problems. Moreover, there may be doubts about indeterminacy of individual rationality of the traditional sort in the interactive setting of games” (*ibidem*). Again, both VFT and TR, even by different approaches of players’ rationality bypass the presupposition that strategic rationality should fits with individual decision theory, and accordingly with individualism. For instance, Bacharach (1995: 1) explicitly stresses that “cooperative reasoning [...] must differ from the type of reasoning standardly assumed in game theory”. In TR, the “unit of agency” is a team and no longer individuals. Players’ mode of reasoning

⁶⁸In rational choice theory, the principle of extensionality entails that “the way options are described should not change your preference between them” (Bacharach, 2001a: 2).

⁶⁹How and why framing entails a violation of the standard conception of the individual decision theory, and especially the expected utility theory is well argued in 2001a.

entirely relies on the team's rationality. In VFT, players' rationality encloses a sense of 'otherness', what it is rational to do for a player intrinsically depends on the others. In a way, in VFT, players integrate in their reasoning that resolving a game is a collective purpose. Players' rationality is no longer a mere individualistic rationality.

Considering the criticism detailed above, and considering the enrichments proposed by Bacharach in both VFT and TR, why identify two concepts of rationality: social rationality and collective rationality?

My view is that players' rationality in VFT and TR is highly different from all the others trend of research investigating the limits of individual decision theory in game theory. First players' preferences (or their reasons for acting) depend on the context, their perceptions, their beliefs, others' perceptions, a cultural determinism, and on self-identities influencing their perceptions; determinants of choice are no longer intrinsic to players. Because players' subjective representations depend on time and place (like for Schelling, 1980 [1960]): "frames may vary both across players and from occasion to occasion" (Bacharach, Bernasconi, 1997: 4) and because "[t]he agent's description matters strategically" (*ibid.*, 5), the outcome of rational plays is necessarily specific, contextual and socially grounded. Both the conceptual and methodological implications of Bacharach's purpose to integrate more realism in the game theoretic approach of players and interpersonal strategic interactions render the players' rationality more complex than in any other behavioural model.

Asserting that players' rationality is collective is less question begging since this term may be used by theorists of TR (e.g. Sugden, Hurley, Gilbert) or those which are aware of this theory and who reflect upon it (e.g. Colman, Hargreaves Heap and Varoufakis). Moreover, Bacharach (1997: 13) asserts that "We-play is rational in the external sense of being efficacious or functional". Team reasoning possesses its own requirements in order to be valid: "cooperative reasoning is *sui generis* and not derived from the standard (individualistic, instrumental) type of reasoning" (Bacharach, 1997: 25).

To conclude, I would like shed light on the fact that because TR shares common determinants with VFT since TR relies on (i) players' frames which are context-dependent, and (ii) a highly comparable conception of players as in VFT, collective rationality may be considered as a subcategory of what I call social rationality. Collective rationality as induced by frames and by players' identity is a specific kind of social rationality. To be more specific social rationality is a premise of collective rationality. That is why, in my opinion, defining players' rationality in VFT, as a social rationality was relevant. It emphasizes that this form of rationality is prior to the collective rationality in Bacharach's work.

5. Conclusion

Throughout this paper I have insisted on the numerous difficulties faced by Bacharach in regard to his conceptual and methodological choices both in VFT and TR. For instance, in VFT two of his major challenges are to draw a game as an objective coordination device and to justify the players' capacity to coordinate. Therefore, he repeatedly requires to *ad hoc* assumptions. In the same manner, his theory of TR enhances considerable troubles mainly materialized into the endogenization problem, and again to justify cooperation he requires to *ad hoc* assumptions. In this perspective Bacharach's achievements do not seem to equal his ambitions. I suggested that this is mainly explained by the fact that Bacharach wants to incorporate subjective determinants in games and players' reasoning, as well as determinants stemming from outside the games and which do not belong to a game theoretical framework. Bacharach seems to take two positions at the same time, thinking of games as a player and as a theorist, without a clear-cut frontier between the "rationality

of the observer” and the “rationality of the observed” (the Schumpeterian dichotomy). This points out that some enrichments of game theory are incompatible with its research program and formalism. It accentuates how game theory is at the same time powerful but limited, since its level of abstraction is too high. Game theory is an untouchable close and coherent system, and blurring some of its frontiers may endanger the models proposed to challenge it. Some paths, and precisely the path taken by Bacharach cannot be reconcilable with the core of standard non-cooperative game theory.

Nonetheless, I would like to highlight Bacharach’s merit in asking worthwhile questions and to underlining the limits of game theory which are rarely discussed and emphasized in such a way. He incorporates social and cultural determinants in a very different way from the current new behavioral economics. For instance, social and “other regarding” preferences (Rabin, 1993; Levine 1998; Bolton, Ockenfels, 2000; Charness, Rabin, 2002; Fehr and Schmidt, 2006; for a selection of the most representative models)⁷⁰ incorporates new arguments in players’ utility functions, without challenging the content of individual standard rationality. The standard conception of economic individuals remains. Players still maximize a *pre-determined* utility function and continue to have the same objective, namely the maximization of their individualistic utility function. The only change is that social preferences “internalize sociality” (Davis, 2011: 89) and that players have “social tastes” (*ibid.*, 85).

Bacharach acknowledges that he could have incorporated other social and cultural determinants in the players’ choice to justify coordination and cooperation, on a comparative ground as “other regarding preferences”:

“Saliency is one sort of potency. But it is not the only one. Among the concepts that are in use as someone deliberates, some might in principle have more influence than others. For example, once someone sees a decision as one involving certain duties, those duties might dominate other considerations” (Bacharach, 2001a: 5-6).

Nevertheless, he would not have enhanced the debates at stakes in both VFT and TR. Besides, Sugden (1993) – among others – stresses that ‘altruism’ and accordingly altruist preferences (a subcategory of “other regarding preference”; see Sobel, 2005) may be inconsistent with standard instrumental rationality (he especially refers to Andreoni [1990]) and may not be sufficient to explain cooperation.

In fact, Bacharach searches to explain individual rationality by “reason within nature”: “reason is a part of nature, so adopting the rational approach does not mean sacrificing the insights of the evolutionary perspective. The method of evolutionary psychology allows us to explain choices by ‘reason within nature’” (Gold, Sugden, in Gold, Sugden, 2006: 7). In fact, “Human beings are endowed with capacities for reasoning that are the product of evolutionary selection; they then apply those capacities to whatever specific decisions problems they face” (*ibid.*, 9). However, quite obviously, his perspective is highly different to the evolutionary game theory⁷¹.

Moreover, even if Bacharach did not stimulate a huge trend of research, some of his intuitions are valuable. The question of framing rarely diffuses into game theory. Although Bacharach opened a few amount of contributions attempting to give a game theoretic content to frames and to explain focal points *via* framing (e.g. Sugden, 1995; Casajus, 1998, 2000; Colman, 1997; Janssen, 2001, 2006; Sugden, Zamarron, 2006). His work participates to a largest though on collective entities, still rare in economics. Both VFT and TR emphasize that players have a sense of otherness and that being successful is a collective enterprise. In a way, in VFT, players integrate in their reasoning that resolving a game is a collective purpose. The ‘other’ is materialized in players’ reasoning since the ‘other’ has perceptions; the ‘other’ is no longer a mere probabilistic event or, shall we say, an

⁷⁰ For a categorization and differences within the label “other regarding preferences” see Sobel (2005).

⁷¹ See Sugden (2001) for a clear-cut representation of the evolutionary game theory.

objective event. In TR, a collective purpose can be reason giving, teams count as a “unit of agency” for a game. Indeed, Bacharach purports to justify that agents can rely on collective entity, and can naturally consider the collective as the relevant “unit of agency” to resolve decision problems. In addition, it is my point of view that Bacharach could be considered as a precursor of the recent “economics of attention”. Both VFT and TR are a matter of attention. The following quotations seem to confirm this viewpoint.

“Part of the motivation for VFT was to reveal the role of rationality in “focussing”” (Bacharach, Bernasconi, 1997: 39; my emphasis).

Salience in VFT is a matter “focus”, or of attention on a success promoting option. In the same manner, when “we frames” are salient in TR, players pay attention to the diagonal of game matrices or “focus their attention” on the characteristics at stakes in “the common problem mechanism”.

“The ‘we’ and ‘I’ ways of thinking about a problem [...] compete for attention; if both recruit an agent’s *attention* – if they are ‘activated’ – they vie for obedience” (Bacharach, 1997: 2; my emphasis).

Finally, it is worthy to note that Bacharach claims that the theory of TR cross the frontiers of non-cooperative game theory. It can necessitate the combination of non-cooperative game theory, cooperative game theory and multi-agent systems (2001b)⁷². From this perspective, it is my opinion that Bacharach offers a new line of research for modelling collective decisions and collective behavior in cognitive economics, in economics of networks and in multi-agent systems. Some of the conceptual and formal aspects of both Bacharach’s VFT and TR could be better suitable within these frameworks than is standard non-cooperative game theory.

Bibliography

Anderlini L. (1990) “Communication, Computability and Common Interest”, *Games and Economic Behavior*, 27: 1-37

Andreoni J. (1990), “Impure Altruism and Donations in Public Good: A Theory of Warm-Glow Giving”, *The Economic Journal*, Vol. 100 (June), p464-477

Aumann R. J., Sorin S. (1989), “Correlated Equilibrium as an Expression of Bayesian Rationality” *Econometrica*, 55(1) : 1-18

Bacharach M. (1976), *Economics and the Theory of Games*, The Macmillan Press LTD,

Bacharach M. (1987), “A theory of rational decision in games”, *Erkenntnis*, 27: 17-55

Bacharach M. (1991), “Games with Concept Sensitive Strategy Spaces”, Working Paper, Institute of Economics and Statistics, University of Oxford

Bacharach M. (1992), “Backward induction and beliefs about oneself”, *Synthese*, 91: 247-284

Bacharach M. (1993), “Variable Universe Game”, in *Frontier of Game Theory*, Ed. Binmore K., Kirman A., Tami P., The MIT Press,

Bacharach M. (1994), “The epistemic structure of a theory of game”, *Theory and Decision*, 37: 7-48

Bacharach M. (1995), “Cooperating without Communicating”, ESRC Economics Beliefs and

⁷² We can found an explicit link between Bacharach’s TR and multi-agent systems in 1999 too.

Behaviour Programme, Discussion Paper No.1

Bacharach M. (1997), "Common Knowledge", published in *The New Palgrave Dictionary of Law and Economics*, Macmillan,

Bacharach M. (1997), "We Equilibria: A Variable Frame Theory of Cooperation", Institute of Economics and Statistics, University of Oxford

Bacharach M. (1998), "Agents with Multiple identities in the Theory of Cooperation", Unpublished Paper

Bacharach M., (1999), "Interactive Team Reasoning: A Contribution to the Theory of Cooperation", *Research in Economics*, 53: 117-147

Bacharach M. (2001a), "Framing and Cognition in Economics: The Bad News and the Goods", Lecture notes, ISER Workshop XIV, July 2001: Cognitive Processes in Economics

Bacharach M. (2001b), "Superagency: Beyond Individualistic Game Theory", Invited Lecture presented at TARK VIII : Eighth Conference on Theoretical Aspects of Rationality and Knowledge, Certosa di Pontignano, University of Siena

Bacharach M. (2006), *Beyond Individual Choice: Team and Frame in Game Theory*, Ed. Gold N., Sugden R., Princeton University Press,

Bacharach M., Hurley S. (1991), "Issues and Advances in the Foundations of Decision Theory", in *Foundations of Decision Theory*, Blackwell publishers,

Bacharach M., Mongin P. (1994), "Epistemic logic and the foundations of game theory", *Theory and Decision*, 36: 1-6

Bacharach M., Bernasconi M. (1997), "The Variable Frame Theory of Focal Points: An experimental Study", *Games and economic Behavior*, 19: 1-45

Bacharach M., Stahl O. (2000), "Variable-Frame Level- n Theory", *Games and Economic Behavior*, 32 (1): 220-246

Bolton G., Ockenfels A. (2000), "ERC: A Theory of Equity, Reciprocity and Competition." *American Economic Review*, 90 (1): 166-193.

Bradsley N., Mehta J., Starmer C., Sugden R. (2010) "Explaining Focal Points: Cognitive Hierarchy Theory Versus Team Reasoning", *The Economic Journal*, 120: 40-79

Brewer M. B., Gardner W. (1996), "Who is the "we"? Levels of collective identity and self-representations", *Journal of Personality and Social Psychology*, 71 (1): 83-93

Casajus A., (1998), "Focal Points in Framed Strategic Forms", *Game and Economic Behavior*, 32: 263-291

Colman A. M. (1997), "Salience and Focusing in Pure Coordination Games", Working Paper, Pub. in *Journal of Economic Methodology*, 4: 61-81

Colman A. M., (2003) "Cooperation, Psychological Game Theory and Limitations of Rationality in

- Social Interaction”, *Behavioral and Brain Sciences*, 26: 139-198
- Crawford V., Haller H. (1990), “Learning How to Cooperate: Optimal Play in Repeated Coordination Games”, *Econometrica*, 58: 571-95
- Davis J. B., (2011) *Individuals and Identity in Economics*, Cambridge University Press, New York
- Farrell J. (1990), “Communication, Coordination and Nash Equilibrium”, *Economic Letters*, 27: 209-14
- Fehr E., Schmid K.M., (1999) “A Theory of Fairness, Competition and Cooperation”, *Quarterly Journal of Economics*, 114 (3): 817-868
- Gauthier D. (1975), “Coordination”, *Dialogue*, 14: 195-221
- Gilbert M. (1989), *On social Facts*, Routledge, London
- Gilbert M. (2000), *Sociality and Responsibility, New Essays in Plural Subject Theory*, Rowman and Littlefield Publishers, Lanham
- Gilbert M. (2003), “The Structure of the Social Atom: Joint Commitment as the Foundation of Human Social Behavior”, in *Socializing Metaphysics: The Nature of Social Reality*, Ed. Friedrich FS, Rowman & Littlefield, Lanham MD, 39-64
- Giocoli N. (2003), *Modeling Rational Agents: From the Interwar Economics to Early Modern Game Theory*, Edward Elgar, Massachusetts, USA
- Gintis H. (2007), “A framework for the unification of the behavioral sciences”, *Behavioral and Brain Sciences*, 30: 1-61
- Gold N., Sugden R. (2007), “Theories of Team Agency”, in *Rationality and Commitment*, Ed. by Peter and Schmidt, Oxford University Press,
- Hargreaves Heap S. P., Varoufakis Y. ([1995] 2004), *Game Theory; A Critical Text*, Second Edition, Routledge, London and New York
- Harsanyi J.C. Selten R. (1988), *A General Theory of Equilibrium Selection in Games*, MIT Press.
- Hédoin C., (2012), “A Framework For Community-Based Salience: Common Knowledge, Common Understanding and Community Membership”, Working Paper, OMI, University of Reims Champagne-Ardenne
- Hurley S. (1989), *Natural Reasons*, New York: Oxford University Press
- Janssen M. C. W., (2001), “Rationalizing Focal Points”, *Theory and Decision*, Vol. 50, p 119-148
- Janssen M. C. W. (2006), “On the Strategic Use of Focal Points in Bargaining Situations”, *Journal of Economic Psychology*, 27: 622-634
- Keren G. (2011), “On the Definition and Possible Underpinnings of Framing Effects: A Brief Review and a Critical Evaluation”, in *Perspectives on Framing*, Ed. Keren, The Society for Judgment and Decision Making Series, Psychology Press, New York, 3-34

- Klaes M. (2008) "Rationality and its Bounds: Re-framing and Social Framing", in *Reasoning, Rationality and Probability*, Ed. Galavotti, Scazzieri, Suppes, CSLI Publication, Chicago University Press
- Kreps D. (1990), "Corporate Culture and Economic Theory", in *Perspective on Positive Political Economy*, Ed. Alt, Shepsle, Cambridge University Press
- Levine D. K. (1998), "Modeling Altruism and Spitefulness in Experiments", *Review of Economic Dynamics*, 1 (3): 593-622.
- Lewis D. (1969), *Convention: A Philosophical Study*, Harvard University Press
- Metha J., Starmer C., Sugden R. (1994), "The Nature of Salience: An Experimental Investigation of Pure Coordination Games", *American Economic Review*, Vol. 84, p 658-673
- Postema, G.J. (2008) "Salience Reasoning", *Topoi*, 27 : 41-55
- Rabin M. (1993), "Incorporating Fairness into Game Theory and Economics", *The American Economic Review*, 83: 1281-1302
- Scazzieri, R. (2008), "Coordination, Context and Patterns of Reasoning", in *Reasoning, Rationality and Probability*, Ed. Galavotti, Scazzieri, Suppes, Chicago University Press, 187-207
- Schelling T. (1980 [1960]), *The Strategy of Conflict*, Harvard University Press,
- Sugden R. (1991), "Rational Choice: A Survey of Contributions from Economics and Philosophy", *The Economic Journal*, 101: 751-785
- Sugden R. (1993), "Thinking as a Team: Towards an Explanation of Non Selfish Behavior", *Social Philosophy and Policy*, 10 (1): 69-89
- Sugden R. (1995), "A Theory of Focal Points", *The Economic Journal*, 105 (May): 533-550
- Sugden R. (2000), "Team Preferences", Working paper, Pub. in *Economics and Philosophy*, 16 (2): 175-204
- Sugden R. (2001), "The Evolutionary Turn in Game Theory", *Journal of Economic Methodology*, 8 (1): 113-130
- Sugden R. (2005), "The Logic of Team Reasoning", in *Teamwork – Multi-Disciplinary Perspectives*, Ed. by Nathalie Gold, Palgrave MacMillan, 181-199
- Sugden R. (2007), "Collective Intentions and Team Agency", *Journal of Philosophy*, 104 (3): 109-137
- Sugden R., Zamarron I. E. (2006), "Finding the Key: The Riddle of Focal Points", *Journal of Economic Psychology*, 27: 609-621
- Tajfel H. (1970), "Experiments in Intergroup Discrimination", *Scientific American Journal*, 223: 96-102
- Tajfel H. (1972) "Social Categorization" *Introduction à la psychologie sociale*, Vol. 1, Ed.

Moscovici S., Larousse, Paris, p272-302

Tajfel H., (1981), *Human Groups and Social Categories: Studies in Social Psychology*, Cambridge University Press, Cambridge

Tajfel H., Turner J. (1985), "An integrative theory of intergroup conflict », in *The Social Psychology of Intergroup Relations*, Ed. By W. Austin and S. Workel, Blackwell Publishers

Tan J. H., Zizzo D.J. (2008), "Groups, Cooperation and Conflict in Games", *The Journal of Socio-Economics*, 37: 1-17

Turner J., (1985), "Social Categorization and the Self-Concept: A Social Cognitive Theory of a Group Behavior", *Advanced in Group Processes: Theory and Research*, Vol. 2, CT: JAI Press, Greenwich, 77-122

Turner et al. (1987), *Rediscovering the Social Group: A Self-Categorization Theory*, Blackwell, Oxford

Tversky A., Kahneman D. (1981), "The Framing of Decisions and the Psychology of Choice", *Sciences, New Series*, 211 (4481): 453-458

Tversky A., Kahneman D. (1986), "Rational Choice and the Framing of Decisions", *Journal of Business*, 59 (4): 251-278

von Neumann J., Morgenstern O. (1944), *The Theory of Games and Economic Behaviour*, Princeton, NJ: Princeton University Press

DOCUMENTS DE TRAVAIL GREDEG PARUS EN 2014
GREDEG Working Papers Released in 2014

- 2014-01** CRISTIANO ANTONELLI & FRANCESCO QUATRARO
The Effects of Biased Technological Changes on Total Factor Productivity: A Rejoinder and New Empirical Evidence
- 2014-02** MICHAËL ASSOUS & ROBERTO LAMPA
Lange's 1938 Model: Dynamics and the "Optimum propensity to consume"
- 2014-03** AGNÈS FESTRÉ & PIERRE GARROUSTE
Do People Stand by their Commitments? Evidence from Classroom Experiments
- 2014-04** SAMIRA DEMARIA & SANDRA RIGOT
IFRS Standards and Insurance Companies: What Stakes for Long-Term Investment? A French Case Explanatory Analysis
- 2014-05** GÉRARD MONDELLO
Splitting Nuclear Parks or Not? The Third Party Liability Role
- 2014-06** FRÉDÉRIC MARTY
Towards an Economics of Convention-based Approach of the European Competition Policy
- 2014-07** GIOVANNI DOSI, GIORGIO FAGIOLO, MAURO NAPOLETANO, ANDREA ROVENTINI & TANIA TREIBICH
Fiscal and Monetary Policies in Complex Evolving Economies
- 2014-08** SÉVERINE BORDERON
La nature devenue projet de compensation écologique
- 2014-09** BENJAMIN MONTMARTIN & MARCOS HERRERA
Internal and External Effects of R&D Subsidies and Fiscal Incentives: Empirical Evidence Using Spatial Dynamic Panel Models
- 2014-10** OLIVIER BRETTE, THOMAS BUHLER, NATHALIE LAZARIC & KEVIN MARECHAL
Reconsidering the Nature and Effects of Habits in Urban Transportation Behaviour
- 2014-11** RAPHAËL CHIAPPINI
Institutional Determinants of Japanese Outward FDI in the Manufacturing Industry
- 2014-12** OLIVIER BRUNO, ANDRÉ CARTAPANIS & ERIC NASICA
Bank Leverage, Financial Fragility and Prudential Regulation
- 2014-13** ANAÏS CARLIN
Consumer Choice Theory and Social Learning
- 2014-14** DJAMILA ELIDRISSI, VALÉRIE HAUCH & SABRINA LOUFRANI-FEDIDA
Compétences relationnelles et entreprises à internationalisation rapide et précoce : une approche multi-niveaux
- 2014-15** FRANCK LÉON
La construction des Business Models des fournisseurs d'Infrastructure as-a-Service : l'étude de deux "purs players" français
- 2014-16** MAËLLE DELLA PERUTA & DOMINIQUE TORRE
Virtual Social Currencies for Unemployed People: Social Networks and Job Market Access
- 2014-17** LAUREN LARROUY
Challenging Standard Non-Cooperative Game Theory? From Bacharach's "Variable Frame Theory" to "Team Reasoning"