We-thinking and 'double-crossing': frames, reasoning and equilibria

Smerilli, Alessandra

University of East Anglia

2 November 2008

WE-THINKING AND 'DOUBLE-CROSSING'
Frames, reasoning and equilibria

Alessandra Smerilli, *University of East Anglia*

*02 November 2008**

"Probability arises from an opposition of contrary chances or causes, by which the mind is not allow'd to fix on either side, but incessantly tost from one to another, and one moment is determin'd to consider an object as existent, and at another moment as the contrary. The imagination or understanding, call it which you please, fluctuates betwixt the opposite views..." (Hume,(1739-1740)[1978]p. 440)

1 Introduction

The idea of we-thinking, or we-reasoning, is increasingly drawing the attention of more and more economists. In its general formulation, it has been proposed by David Hodgson (1967), Donald Regan (1980), Margaret Gilbert (1989), Susan Hurley (1989), Raimo Tuomela (1995), and Martin Hollis (1998). Within this body of literature, Robert Sugden (1993, 2000, 2003) and Michael Bacharach (1995, 1997, 1999, 2006) have developed analytical frameworks from an economic point of view. The main claim of scholars that analyze we-thinking is that it can be endorsed by people when they face a decision problem. In fact, experimental evidence shows that, especially in some kind of games, such as coordination games, people do endorse we-thinking[1].However, the way in which we-thinking arises and how it brings people to behave in a particular way in games is a matter that requires further investigation.

The two main contributors are Bacharach and Sugden, and they approach the topic in two different ways. Sugden's aim is to show that we-reasoning is

*I wish to thank Federica Alberti, Luigino Bruni, Benedetto Gui, Fabrizio Panebianco, Vittorio Pelligra, Lorenzo Sacconi, Robert Sugden, and Stefano Zamagni for their valuable comments and advice, although responsibility for any errors remains mine alone.

[1]See Tan and Zizzo (2008) for a review of experiments.

a consistent and logical way of thinking, but he does not face the problem of how we-reasoning can arise. He gives some intuitions only about a psychological background based on Smith's analysis of correspondence of sentiments. Bacharach's theory is based on frames and his never reached aim (because of his death) was to explain we-thinking in terms of Variable Frame Theory (Bacharch 1993). But, as we shall see later, some of his intuitions conflict with the logical analysis he proposes. His logical analysis was developed by Zizzo and Tan (2003, 2008): They introduce a 'game harmony' measure, as a proxy of the extent of cooperation or conflict in games. Game harmony could represent a step forward towards the endogenization of the probability that we-thinking could arise in a decision problem.

In the present paper, I take a different approach to the way in which we-thinking works. Based on a not fully developed intuition of Bacharach's, that does not entirely fit into his logical analysis, i.e. the 'double-crossing' problem in Prisoners' Dilemma (PD) game, I propose a framework in which a person is allowed to have both I-thoughts, when she is we-reasoning, and we-concepts, when she is I-reasoning, and develop my analysis in terms of equilibrium concepts.

2 Bacharach's theory of we-thinking

"The answers to fundamental questions about coordination and cooperation... lie in the agent's conception not of the objects of choice, nor of the consequences, but of herself and of the agents with whom she is interacting"(Bacharach, M., N. Gold and R. Sugden 2006, p. 70). This sentence is the starting point of Bacharach's analysis of we-reasoning.[2] We-reasoning is

[2]It is interesting to see how and when Bacharach developed the idea of we-thinking. He started by building the Variable Frame theory (Bacharach 1993), in parallel he was developing a theory of cooperation. In 1995 he introduced the category of 'fellow member reasoner': "Someone who is a member of a natural type T and chooses a certain strategy if she is sufficiently sure that her interactants are also member of T" (- 1995, p.1). In this context he tries to link T-membership to variable frame theory and, at for the first time he introduces the 'we' category: "The present paper has made type T membership an issue which type T members think about, and nuanced their capacity to recognize it. An alternative development would make T membership a variable element in players' frames in the sense of variable frame theory: 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" (p.17). In 1997 Bacharach formally introduces we-thinking, but, at the same time, he claims that we-reasoning is different from team-thinking: "the essence

seen as a powerful 'mechanism' (in Bacharach's words) for solving puzzles about cooperation and coordination in game theory (i.e. games like Hi-Lo and Prisoner Dilemma). More in general, in his work Bacharach tries to demonstrate, by showing some evidence,[3]that we-reasoning is a valid mode of reasoning and people do endorse it.

His theory is based on frames: if we-frame comes to mind, the subject will group identify and then she will start to we-reason. A frame can be defined as a set of concepts that an agent uses when she is thinking about a decision problem. It cannot be chosen, and how it comes to mind is a psychological process: "Her frame stands to her thoughts as a set of axes does to a graph; it circumscribes the thoughts that are logically possible for her (not ever but at the time). In a decision problem, everything is up for framing...also up for framing are her coplayers, and herself". (ib. p. 69). In Bacharach's framework, then, a person may start to we-reason only if she has 'we' concepts in her frame: in other words, a person firstly recognizes the we-perspective, and then endorses it.

The 'framing' perspective is not the only way in which the theory of we-thinking has been proposed, however. Robert Sugden, for example, has developed a different framework for looking at the problem. In his framework the central concept is the 'common reason to believe'[4]: people who group identify are not committed to reason as a team unless there is a common reason to believe that other agents are doing the same. The psychological

---

of team thinking is that each player participates in the group-best profile in conditions of common knowledge that they form a 'team' – that is, that they all act in this way" (p.13). The theory of 'team-reasoning' he has in mind was proposed by Robert Sugden in 1993. This was a seminal paper in Sugden's research project on we-thinking. In fact later versions of his theory (see Sugden 2000, 2003) do not assume the participation of all subject to participate and the common knowledge hypothesis. The first published paper in which Bacharach formalizes his theory is the 1999's article about 'interactive team reasoning'. In it Bacharach introduces some elements that we can find in the book, such as, group identification, team reasoning as the effect of group identification, unreliable team interaction, that in the book becomes cirscumspect team reasoning, etc. Between the '99 article and the book we may find some lecture notes, in which the concepts of agency and 'superagency' begin to appear. The book represents an (incomplete, because of his death) attempt to build a complete theory of we-thinking.

[3]Bacharach claims that there are five kinds of evidence: logical, introspective, evolutionary, transcendental and experimental (see Bacharach 2006, pp.145-146).

[4]Common reason to believe is defined as follows: "there is common reason to believe to a preposition p in a set of individual T if: (i) for all individuals i in T, i has reason to believe p (i.e. p can be inferred from propositions that she accepts as true n.d.a.); (ii) for all individuals I and j in T, I has reason to believe that j has reason o believe p; (iii) for all individuals I, j, and k in T, I has reason to believe, that j has reason to believe that k has reason to believe p; and so on" (Gold and Sugden 2008, p. 302).

side of group identification in Sugden's theory might be found in his analysis of Smith's 'correspondence of sentiments' (Sugden, R. 2005): 'fellow-feeling', could be seen as the source of group identification. In sum, in Bacharach's framework if people group identify they automatically start to reason like a team, whereas in Sugden's theory people may group identify, but team reasoning does not follow automatically.

Sugden's framework can be summarized as in figure 1.

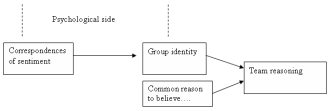

Figure 1: Sugden's framework

Bacharach's aim is to explain situations in which some people may 'we'-reason and some other may not. In order to model the previous intuition, he assumes that a frame 'we' can come to mind with a probability 'ω', which represents the probability for a subject to group-identify. The probability ω is common knowledge[5], in fact: "in coming to frame a situation as a problem 'for us', an individual also gain some sense of how likely it is that another individual would frame it in the same way" (Bacharach, Gold and Sugden 2006 p. 163). A context in which some people may group-identify and some may not is seen by Bacharach as an unreliable coordination context, and team reasoning in this context is called circumspect team reasoning.

In Bacharach's theory group identification is a framing phenomenon that determines choices by "changing the logic by which people reason about what to do" (ib). If, by reasoning in the individual standard mode (I-reasoning), an agent looks at a decision problem by thinking what it would be the best for her to do, when there is group identification, the agent will think: "What would the best be for us to do?". Basically then, "Somebody 'team reasons' if she works out the best feasible combination of actions for all the members of her team, then does her part in it"(Bacharach 2006 p.121).

Sugden describes team reasoning in a similar way: "The idea is that, in relation to a specific decision problem, an individual may conceive of herself as a member of a group or team, and conceive of the decision problem, not as a problem for her but as a problem for the team. In other words, the

[5]In a previous work (1999), Bacharach has developed a more formalized model, in which each agent can participate or lapse in a team and everyone, before choosing, receive a signal knowing the joint probability distribution of this signal and agent's state (i.e. an agent's signal includes her participation state).

individual frames the problem, not as 'What should I do?', but as 'What should we do?'"(2000, pp. 182-183).

Trying to go deeper into the matter, Bacharach aims to explain how a frame, in particular a we-frame, may come to mind. In his earlier works (Bacharach, M. 1997, 1999) he proposes that the possibility of team reasoning is related to having 'scope for cooperation' and to the 'harmony of interests'.

In his last book he suggests the (strong) Interdependence Hypothesis, that roughly states: perceived interdependence prompts group identification. The perception of interdependence between two agents in a game is given by three factors:

- common interest (the agents have common interest in s* over s, if both prefer s* to s, where s*, s are possible state of affairs, or, in a game, possible outcomes)
- copower (nobody can reach s* alone, but both can together)
- standard solution (basically Nash equilibrium that may realise s).

Bacharach says that if the two agents have common interest in s* over s and copower over it, and the standard solution of the game contains outcomes in s (this is the case of PD), then people are stimulated to group identify. In other words, if s is a possible outcome of the individual rationality, and s* is strictly Pareto-preferred by both the agents, and they have common power to bring about s*, then group identification is stimulated. Or, if the outcome that can be reached by an individual way of reasoning is Pareto-dominated by another outcome achievable only by thinking as a group, there is space for group identification.

We may summarize Bacharach' framework in the scheme in figure 2.

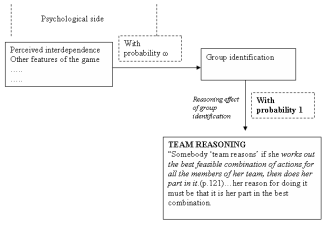

Figure 2: Bacharach's framework

The main Bacharach's purpose is to explain cooperation, seen as a successful group activity (see 2006 p. 69), and the core mechanism for doing

that comprehends 'framing', 'common purpose', and 'cooperation': "(i) we frame ourselves as members of groups; (ii) ... perceived agreement of individual goals among a set of individuals favours framing as members of a group with this common goal; (iii) the group framing tends to issue in efficient cooperation for the group goal" (p.90). People then cooperate because they group identify.

3 We-thinking and Variable Frame Theory

Bacharch's (never reached) aim was to explain we-reasoning in terms of Variable Frame Theory (VFT), which he had developed in a earlier stage of his investigation[6]. Concisely, in VFT a player can intentionally choose an object, or an action, if she has a way of thinking about that object or that action, i.e. he has a frame. Frames can be more or less salient or available, depending on a probability measure on them. A decision rule in VFT is "a mapping from frames to options induced by those frames" (Bacharach, M. 2001a), and an equilibrium for symmetrical games is defined as follows: "the pair (δ,δ) is a variable frame equilibrium if, for each Frame F, the option δ(F) is subjectively best from the perspective of F against δ as perceived in F" (ib.). In other words the decision rule δ(F) has to be the best reply against δ.

The intersection between VFT and we-thinking would have been called by Bacharach 'Variable agency theory' (Bacharach 2006, p.59). However, he could not complete the description of 'we'-reasoning in terms of VFT. In fact, there are at least two problems to solve, in order to complete Bacharach's theory: one is related to the way he conceives the 'we' frame, the other is the endogenization of ω.

In Bacharach's circumspect team reasoning, as I have said before, if people group identify, then the we-frame comes to their mind and they start to we-reason. It seems as though in Bacharach's framing theory there are two aspects that are deeply linked: in framing a situation, the first step is to recognize a frame, that is coming to see it; the second step is endorsing that frame, i.e. reasoning as the frame allows you to do. In Bacharach's theory group identification means not only endorsing a particular way of reasoning, but also coming to see it. The 'compression' between the two aspects of framing is due to the VFT. In it in fact, changing frame does not mean to change the way of reasoning and the decision problem for a subject is fully determined by the interplay of his frame and the objective world.

[6] See Bacharach 1993, 2001

Because of this 'compression', Bacharach in his theory of we-thinking cannot allow people to use more than one frame at a time. In a certain sense, as it has been noticed by Gold and Sugden (Bacharach, Gold and Sugden 2006), in we frame people become committed to we-reason: "In the theory of team reasoning, an individual who reasons in the 'we' frame is aware of the 'I' frame too (as one of that other players might use) but acknowledge only 'we' reasons. It seems that group identification involves something more that framing in the sense of variable frame theory: the group-identifier does not merely become aware of group concepts, she also becomes committed to the priority of group concepts over individual ones" (p.199).

The fact that people cannot use more that one frame at time and they cannot reason about frames whilst thinking of which frame to endorse, prevents Bacharach to develop analytically one of his intuitions. In fact, taking the most famous game in terms of cooperation, the PD game, as an example, Bacharach says: "In a Prisoner's Dilemma, players might see only, or most powerfully, the feature of common interest and reciprocal dependence which lie in the payoffs on the main diagonal" (p.86). If this happens, players do cooperate. But, it might be the case that "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 in the expectation that she will play C. This perceived feature might inhibit group identification" (ib).

Here Bacharach seems to have in mind some psychological process which inhibits group identity which is not quite represented by his own concept of interdependence – the idea of 'double-crossing'. The reason this idea doesn't fit his framework is that double-crossing is the incentive to act on individual reasoning when one believe the other is acting on team reasoning. This requires that the player uses both frames at the same time, while thinking about which one to use. Or, in other words, a player, in order to recognize the 'double-cross' threat, should be allowed to imagine himself in a we-frame, and then deliberating to cooperate, but at the same time he should use the I-frame by thinking that the other player would take advantage of her. In the first player's conjecture, the other player too should use the we-frame in order to think that the first player could choose to cooperate, and, at the same time, she should use I-frame in order to think how 'double cross' the first player.

In the theory of we-thinking the way in which a person reasons (I-mode or we-mode) is a consequence of the perceived frame. So, if a person is in we-frame she cannot reason in the standard theoretical mode, and then she cannot 'see' the double-crossing threat. She may switch from I-mode of

reasoning to we-reasoning (if we-frame comes to mind), or not. Bacharach, then, does not seem to take into account the possibility that once we are in the we-frame, we may switch to I-mode of reasoning, or better, he allows the possibility of switching frame, but does not allow a person to be able to visualize switching frames. And this is why he cannot represent his 'double-crossing' intuition. It seems that when the "we" frame is perceived, it is also perceived as the correct frame or dominant frame, so that once a person sees the world this way she cannot visualize going back to seeing it the other (compare illusions, myths, lies – 'the scales fell from my eyes').

In effect, in one of his unpublished papers (Bacharach, M. 1997), Bacharach allowed for the possibility of the existence of three frames: the I frame, the We frame and the 'S' (superordinate) frame. We and I are called simple frames: "players in them begin their reasoning with the two basic conceptualization of the situation, as 'what shall we do?' problem and 'what shall I do?' problem respectively" (p. 5). A S frame is active when someone manages "during deliberation to see the problem from both the we and the I/she perspectives"(p.14). Although Bacharach allows for the existence of S, based on psychological attainments, he states that we and I perspectives cannot be held simultaneously: "Although we can switch self-identities rather easily, we appear to be unable to inhabit more than one at a time" (p.15). This seems to be a contradiction, and in fact this assumption leads Bacharach to introduce a solution concept that imposes, in his words "a stringent requirement" (p.21). He assumes that I thoughts in S frame generate a personal evaluation, whereas we thoughts generate a group evaluation. The solution concept in the model roughly states that the cooperative option is chosen by a player in S if it is the best (against the mix of the coplayer's generated by his varying frame –p.21 -) in group evaluation and not worse than the other option in personal evaluation.

The S-frame intuition of 1997 unpublished paper, however, disappeared in subsequent pieces of work, but also the hypothesis that agent can 'vacillate' between the two frames does not appear in the most recent Bacharach's effort: the book.

Later on, in developing the VFT Bacharach faces the issue of integrability of frames. He says that normally frames are integrable: "It is easy to integrate frames which consist of classifiers such as shape, colour and position: we can easily see a mark as a triangle, as a blue triangle, as a blue triangle on the left,... on the other hand... a person can see the marks as letters and as geometric shapes, but not at the same time – you can't integrate these two perception" (2001 a, p.6). There exist frames, then, that are non-integrable. 'I' and 'we' frames appear non integrable in Bacharach's

words, and when this happens, "the agent may find herself vacillating between the judgments that she should do" (ib.). In spite of these attempts, then, 'double-cross' threat has not been enclosed in Bacharach's analytical framework.

I shall suggest later that it is possible to take into account what Bacharach called 'personal' and 'group' evaluation, by reasoning in terms of deviation from an equilibrium and not in terms of frames. Or better, it is possible to do that, if we separate the two aspects of framing: how a frame might come to mind and how people endorse a particular frame when she sees more than one frame. In this way it would become possible to represent double-cross intuition.

The second unsolved problem is the endogenization of ω. Bacharach tries to endogenize ω, because he sees that the fact that ω is exogenous represents a lacuna in his theory. We may see a clear evolution in his thoughts along the years: in 1997 and 1999 papers he sees ω as a function of the gain from cooperation and the harmony of interests: "To endogenize ω, and other feature of 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. The laboratory evidence is promising, as it suggests that group identification may be induced by the 'common problem' mechanism'. In addition, it is plausible that ω may be an increasing function of certain quantitative features of the payoff structure, such as 'scope for co-operation' and 'harmony of interest'" (1999, p.144). In 2001 he turns to other two features: "We need a link from the game parameters to the onset of group identification. Two of the classic favouring conditions are commonness of interests and commonness of predicament" (2001 b, p.8). In his book he goes deeper into the matter, as we have seen, by stating the interdependence hypothesis. But he does not complete the work: "It may also be that there is a positive relationship between salience and effectiveness: when a feature tending to promote self-identity is highly salient, then if and when it is noticed it is also highly effective. These are empirical speculations; their investigation will be an important part of the future development of the theory of group action" (2006, p. 87).

A step forward on this topic has been made by Tan and Zizzo (Tan, J. and D. Zizzo 2008): in their paper there is an attempt to investigate the relationship between harmony of interests ('game harmony' for them), group identification and cooperation. They claim that game harmony is a good measure of the extent of cooperation or conflict in games. However, in experiments and in real life, we often observe what they call excess conflict or cooperation relative to theoretical predictions. In their framework, it is

group identification that can explain excess in-group cooperation and excess out-group conflict. They also believe that the so called 'perceived harmony' can capture the effect of team reasoning on cooperation. This is a theoretical construct and it has not been tested yet. At the same time 'perceived harmony' cannot be inferred by the payoffs of the game, because it is a sort of 'experimental' measure (similarity index in experiments – how subjects think a game is similar to a pure coordination game or to a zero-sum game - is a proxy for perceived harmony). Anyway, the simple game harmony (not the perceived one) is the best existent proxy for what Bacharach has called 'the harmony of interest', and it is entirely derived from the payoffs of the game. Game harmony, defined as "a generic property describing how harmonious or disharmonious the interests of players are, as embodied in the payoffs" (Tan and Zizzo 2008, p. 3), is based on the correlation coefficient between payoff pairs. This measure can be a solution of Bacharach's problem of endogenization of ω. However, as we shall see in section n. 6 it presents some problems. In the next section, I will show a possible way to represent Bacharach's intuition of 'double-cross' problem.

4 Representing the 'double-cross intuition': reasoning in terms of deviations from equilibrium

Looking at the situation and at the way of reasoning from another perspective may help us to enrich the analysis. In what follows, I shall reason in terms of individual and collective rationality as two alternative ways of approaching a decision problem, and in particular I shall focus on reasons for deviating from an equilibrium.

First of all, I suppose that the group utility function of a combination of actions (a1, a2), when the players are P1 and P2, is U(a1, a2) = (u1(a1, a2) + u2(a1, a2))/2[7], where u1(a1, a2) and u2(a1, a2) are the player's payoffs. A player who team reasons, first computes which is the best profile for the group[8], and then he does its part in it. A player who 'I'-reasons follows the standard theoretic predictions of game theory.

It is possible to classify games in terms of reasoning on deviations. The basic idea is that a person may reason in the standard I-mode, or in we-mode, but she may have both frames (I and we) in mind (perhaps not at the same time, if the non-integrability hypothesis is correct, but vacillating between

[7] This formulation is the most used one in literature.

[8] The best profile is calculated allowing for 'unreliable' contexts, and then taking into account the possibility that not everybody will group-indentify.

them). In standard game theory an equilibrium is defined as a combination of actions in which no player has anything to gain by changing unilaterally her own strategy. In we-reasoning theory, an equilibrium is instead defined as a combination of actions in which the whole group cannot gain anything by switching from this combination to another. Deviation is seen then as a test for the existence of an equilibrium, no matter if I or we-equilibrium. In what follows, I shall simply test games in search for equilibria that hold from both I and we point of view.

Table 1: game A

	L	R
H	***3,3***	4,1
L	1,4	2,2

Take, for example, the game A (tab. 1). The combination of actions (H.L) is a Nash equilibrium. In it, in fact neither row player nor column player has reason to unilaterally deviate from that combination of actions. But the same combination is also a we-equilibrium: as a group both players cannot do anything better by switching to another combination[9].

Table 2: game B

	L	R
H	***2,2***	3,0
L	0,3	***2,2***

Game B shows a unique Nash equilibrium, (H,L) and two we-equilibria, (H,L) and (L,R), but only the (H,L) combination is an equilibrium at the same time for I and we-reasoners.

Table 3: game C

	L	R
H	**3,3**	1,1
L	1,1	***2,2***

Game C is an Hi-Lo game, and as it is well known, it has two Nash equilibria, i.e. (H,L) and (L,R), but only one we-equilibrium that is (H,L).

[9]The utility U for the group is 3 in (H,L) combination, 2.5 in both (H,R) and (L,L), and 2 in (L,R).

Table 4: game D

	L	R
H	***3,3***	1,4
L	4,1	**2,2**

Prisoner's Dilemma game, instead, has one Nash equilibrium (L,R) and one we-equilibrium (L,R), but they do not coincide.

If an equilibrium live through both, I and we, deviation tests, it is strong, in the sense that it allows for the existence of both ways of reasoning. At the same time such equilibrium could be seen as a refinement when more than one equilibrium exists. In game B, for example, there are two we-equilibria, but if we allow players to see the game endorsing both I and we concepts, this could help them to recognize that (H,L) equilibrium is the prominent one, because it passes both deviation tests. In this case, having an I thought helps we-reasoners to select an equilibrium. But the opposite can happen as it is in the Hi-Lo game, where there are two Nash equilibria and we-thoughts can help I-reasoners to choose the (H,L) equilibrium.

The double test for deviation could also be seen in terms of deliberations, and not only as a method for testing the existence of an equilibrium. The scheme in fig. n. 7 represents a possible way to classify the previous games in terms of deliberation.

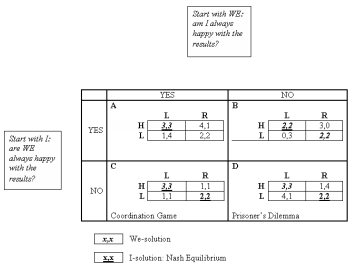

Figure 3: Reasoning about deviations and deliberations

Take for example the game A: in this game, if I start to reason in the standard I-mode we, as a group will be happy with the result (H,L), i.e. we won't want to deviate jointly from the I-reasoning 'solution'. On the opposite, if I group identify, and then I look for the best solution for the group, I as an individual will be happy with the result, i.e. I won't want to deviate unilaterally from the we-reasoning solution. So, in this game, the same result

will be reached, independently of the particular way of reasoning. We may say that I or we-reasoning are observationally indifferent because they give the same result in terms of choice. Looking at this game from Bacharach's perspective (at least my interpretation of it), I notice that this game does not possess the features that prompt we-reasoning (because the individual solution is not Pareto-dominated). This example illustrates how Bacharach's perspective does not appraise we solutions from the I perspective, but only the other way round. One may represent Bacharach's criterion as the answer to the following question: is there an I-solution which we are unhappy with? This will be represented by the lower part of the scheme (and in fact coordination games and prisoner's dilemma games are considered by Bacharach as games in which we-reasoning can arise).

But there could be different situations. Let us look at the game B: in this case, if I start with I-mode, there will be a unique Nash equilibrium (H,L), that is also one of the two possible (and indifferent) we-solutions. If I start with I-mode, we shall then be happy with the result. But, if we group identify and we-reason, if we-reasoning gets us to (H, L), I am happy. If it gets us to (L, R), I am unhappy. So the end of the story is the outcome (H, L), either by we-reasoning or by I-reasoning. This result is observationally equivalent to I-reasoning but not to we-reasoning, because the latter allows (L, R). For the same reasons as before, the game B belongs to the border line in Bacharach's view (I and we solution are the same in terms of payoff).

Game C, instead (the Hi-Lo game), will prompt we-reasoning: if I start with I-mode we won't be happy, because I cannot choose between the two NE, only one of which is also a we-solution; but if we group identify I will be happy with the result.

The last game, the Prisoner's Dilemma game, is the most interesting: if I start with I-reasoning, we won't be happy (the NE is Pareto-dominated by the we solution). But if we group identify the we solution is not good for me (I would be better off by playing the other strategy). In this case there can be a continuous switching from a frame to another. In fact, in the experiment on the PD game, we observe a rate of cooperation of about 50% (see Sally 1995). Following Bacharach's theory, the PD, as we have seen, is the typical game that can lead to we-reason, by prompting the 'interdependence hypothesis'. In the framework I have presented, instead, the 'double-cross' intuition is taken into account, and this generate a perpetual shift, and then we-reasoning is only one of the two equally possible solutions. It is plausible that in cases like this, the salience of frames will play a key-role in the selection of the solution of the game.

This way of looking at a decision problem does not tell us which frame

is more likely to appear. But, if a frame comes to mind, within this classification, we may see, depending on the kind of game the subject is facing, if the frame will be stable or not. However, 'being happy with the results' sometimes depends on which solution is selected. In game B, starting with WE, I may or may not be happy. In game C, starting with I, WE may or may not be happy. We must read, then, the line 'yes' in the classification of fig. 3, as the situation in which agents are not willing to deviate whatever solution will be selected.

5 Discussion

As I said before, my analysis tries to expand an intuition not fully developed by Bacharach. Tan and Zizzo's proposal is also a way of developing Bacharach's analysis. It is possible to make a comparison in terms of predictions, between Bacharach's view, based on strong interdependence hypothesis, Tan and Zizzo's 'game harmony' and my own proposal. Let us see the scheme in table 5.

Table 5:

GAME	BACHARACH'S PREDICTION	GAME HARMONY	PREVAILING WAY OF REASONING
A	I-reasoning	- 0,8	I or We
B	Border Line	-0,9	I
C	We-reasoning	1	We
D	We-reasoning	-0,8	Conflict (perpetual shift)

Game A does not fit the interdependence hypothesis, and then, in Bacharach's framework it prompts I-reasoning. It has a negative harmony, and then in Tan and Zizzo's framework it is seen as a conflictual game. In my framework, both I and we-reasoning are possible and both of them lead to the same solution of the game. Game B, instead, belongs to the borderline in terms of interdependence, because the I solution of the game is not strictly Pareto-dominated, but equal to the we-solution. It is seen as conflictual game in terms of harmony and the prevailing way of reasoning it prompts is I. Game C, the Hi-lo game has the same prediction from every point of view. In contrast, the PD game could bring an agent to team reason by following Bacharach; it is conflictual in terms of game harmony (and then does not lead to cooperative option); and it can lead to a perpetual shift in term of

reasoning, from my point of view.

Games A and D have the same harmony (- 0.8), but they are different in terms of thinking about deviations (yes/yes for A and no/no for D). At the same time, by slightly changing the payoffs of Game A, harmony will change but not the way of reasoning.

Table 6:

	L	R
H	***4,4***	3,1
L	1,3	***2,2***

The game in table 6 belongs to category A (yes/yes) but now the game harmony has become positive: it is 0.2. Then, the I-solution (Nash equilibrium) remains the same, like the we-solution. But game harmony switches from a negative value to a positive one, that is from conflictual to cooperative prediction.

Table 7:

	L	R	S
H	***2,2***	3,0	2,1
L	0,3	***2,2***	2,1
T	1,2	1,2	1,1

By adding row T and column S to game B (as in tab. 7), the way of reasoning does not change, but harmony does change: from – 1 to -0,7. The way of reasoning does not change because the added row and column are dominated from both, I and we point of view. Similarly, by adding row T and column S to the game C (Hi-Lo game, as in tab. 8), it remains the same type of game in terms of way of reasoning (no/yes), but harmony changes (it becomes 0,16, instead of 1).

Table 8:

	L	R	S
H	**3,3**	1,1	0,2
L	1,1	***2,2***	2,0
T	2,0	0,2	0,0

The previous examples tell us that perhaps game harmony measure needs

to be refined. Or better, game harmony is not a strategic way to look at a game, and then we cannot ask to this measure to embody strategic reason. And in some way this measure does respect Bacharach's approach to frames. In fact, in Bacharach's analysis coming to see a frame is a matter of salience, and games with a high game harmony do make we-solution, and then we-frame more salient. However, in some ways game harmony measure and my proposal are linked. In fact, a step forward towards a theory of choice that takes in account both I and we modes of reasoning might be a comparison between incentives to deviate from an equilibrium. Take for example the following games (table 9), that represent PD games with increasingly 'game harmony' measures.

	L	R
H	***3,3***	0,5
L	5,0	**2,2**

	L	R
H	***3,3***	1,4
L	4,1	**2,2**

	L	R
H	***3,3***	1.9,3.9
L	3.1,1.9	**2,2**

Table 9: games with increasing degrees of harmony

The incentive for deviating from a we-equilibrium is measured as the difference between the payoff in (L,L) and the one in (H,L) for the row player[10]It means 2 for the first game, the less harmonious, 1 for the second and 0.1 for the third, the most harmonious. In the third game, thus, we-reasoning is more likely to appear than in the first one. Although further investigation is needed, this seems to be a promising approach to we-thinking theory and more in general to decision theory.

[10]The game is symmetric and then the result will be the same if we calculate the incentive for the column player.

6 Conclusions

In this paper I have analysed Bacharach's theory of we-thinking. In particular I focused on his attempt to formalize we-thinking in terms of Variable Frame theory. I found that two main problems arise trying to do it. One is linked to the endogenization of ω, the probability that a person may group-identify in a determinate situation. The other one is the impossibility, in Bacharach's framework, to use more than one frame at the same time. This problem prevents Bacharach from representing his 'double-crossing' intuition in the Prisoner Dilemma game, because it requires an agent to have I-frame in mind when he is we-thinking. I have proposed a way in which the 'double-cross' intuition may be taken into account: reasoning about deviation from equilibrium, where equilibrium is seen both from an I and from a we point of view. This requires that an agent might reason using more than one frame at time, and it is not a psychological theory. What I have presented it is only a first step based on an equilibrium concept. Steps forward might be to verify if people really endorse this way of reasoning, and to formalize what happens if some people group-identify and some others do not in this framework, that is, to allow for the existence of unreliable contexts.

References

[1] BACHARACH, M. (1993): "Variable Universe Games," in Frontiers of Game Theory, ed. by K. Binmore, A. Kirman, and P. Tani. Massachusetts: MIT Press.

[2] BACHARACH, M. (1995): "Co-Operating without Communicating," London.

[3] BACHARACH, M. (1997): ""We" Equilibria: A Variable Frame Theory of Cooperation," Oxford: Institute of Economics and Statistics, University of Oxford, 30.

[4] BACHARACH, M. (1999): "Interactive Team Reasoning: A Contribution to the Theory of Cooperation," Research in Economics, 53, 30.

[5] BACHARACH, M. (2001): "Framing and Cognition in Economics: The Bad News and the Good," ISER Workshop, Cognitive Processes in Economics.

[6] BACHARACH, M., N. GOLD, and R. SUGDEN (2006): Beyond Individual Choice. Princeton University Press.

[7] GILBERT, M. (1989): On Social Facts. Routledge.

[8] GOLD, N., and R. SUGDEN (2008): "Theories of Team Agency," in Rationality and Commitment, ed. by P. Di Fabienne, and S. H.: Oxford University Press.

[9] HODGSON, D. H. (1967): Consequences of Utilitarianism. Oxford: Clarendon Press.

[10] HOLLIS, M. (1998): Trust within Reason. Cambridge: Cambridge University Press.

[11] HUME, D. ([1739] 1978): A Treatise of Human Nature. Oxford: Oxford University Press.

[12] HURLEY, S. (1989): Natural Reasons. Oxford: Oxford University Press.

[13] REGAN, D. (1980): Utilitarianism and Cooperation. Oxford: Clarendon Press.

[14] SALLY, D. (1995): "Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiment from 1958 to 1992," Rationality and Society, 7, 58-92.

[15] SUGDEN, R. (1993): "Thinking as a Team: Toward an Explanation of Nonselfish Behavior," Social Philosophy and Policy, 10, 69-89.

[16] SUGDEN, R. (2000): "Team Preferences," Economics and Philosophy, 16, 175-204.

[17] SUGDEN, R. (2003): "The Logic of Team Reasoning," Philosophical explorations, 16, 165-181.

[18] SUGDEN, R.(2005): "Fellow-Feeling," in Economics and Social Interactions, ed. by B. Gui, and R. Sugden: Cambridge University Press.

[19] TAN, J., and D. ZIZZO (2008): "Groups, Cooperation and Conflict in Games," The Journal of Socio-Economics, 37, 1-17.

[20] TUOMELA, R. (1995): The Importance of Us: A Philosophical Study of Basic Social Notions. Stanford University Press.

[21] ZIZZO, D., and J. TAN (2003): "Game Harmony as a Predictor of Cooperation in 2 X 2 Games: An Experimental Study," Oxford: Department of Economics, University of Oxford.