

## Bounded best-response and collective-optimality reasoning in coordination games

Marco Faillo

Alessandra Smerilli

Robert Sugden

22 March 2017

### Abstract

We investigate experimentally the conditions under which bounded best-response and collective-optimality reasoning are used in coordination games. Using level- $k$  and team reasoning theories as exemplars, we study games with three pure-strategy equilibria, two of which are mutually isomorphic. The third is always team-optimal, but whether it is predicted by level- $k$  theory differs across games. We find that collective-optimality reasoning is facilitated if the collectively optimal equilibrium gives more equal payoffs than the others, and is inhibited if that equilibrium is Pareto-dominated by the others, considered separately. We suggest that coordination cannot be explained by a single theory.

**Keywords:** team reasoning, level- $k$  theory, coordination games

**JEL codes:** C7, C9

\*Faillo: Department of Economics and Management, University of Trento, 38122, Italy; email: [marco.faillo@unitn.it](mailto:marco.faillo@unitn.it). Smerilli: PFSE-Auxilium, via Cremolino 141, Rome; email: [asmerilli@pfse-auxilium.org](mailto:asmerilli@pfse-auxilium.org). Sugden (corresponding author): School of Economics and Centre for Behavioural and Experimental Social Science, University of East Anglia, Norwich, UK, NR4 7TJ; email: [r.sugden@uea.ac.uk](mailto:r.sugden@uea.ac.uk).

**Acknowledgements:** The research reported in this paper was funded by the Italian Ministry of Education, University and Research under the national research project (PRIN) no. 20085BHY5T and by the University of Trento. We thank an Editor and two anonymous referees for constructive comments.

One-shot coordination games are interesting, not only as models of a significant class of real-world interactions, but also as posing a puzzle for game theory. Applied to any such game, the classical assumptions of best-response reasoning and common knowledge of rationality identify a set of two or more Nash equilibria, but do not select any one of these equilibria as ‘the’ solution. In many cases, however, real players of these games are remarkably successful at coordinating their behaviour. The puzzle is to explain how they coordinate, and on which equilibrium.

In their attempts to explain behaviour in one-shot coordination games, game theorists have given particular attention to two distinct modelling strategies. The *collective optimality* approach is exemplified by the principle of *payoff dominance* in Harsanyi and Selten’s (1988) theory of equilibrium selection and by the theory of *team reasoning* (Sugden, 1993; Bacharach, 1999, 2006). This approach assumes that rational players appraise strategy profiles from the viewpoint of the players collectively. Leaving aside complications resulting from isomorphisms (which will be discussed later), the payoff dominance principle selects the Nash equilibrium profile (if one exists) that Pareto-dominates all other such equilibria. Team reasoning selects the strategy profile (if one exists) that is uniquely optimal for the players collectively. The *bounded best response* approach is exemplified by *level-k theory* (Stahl and Wilson, 1994; Nagel, 1995; Costa-Gomes, Crawford and Broseta, 2001) and *cognitive hierarchy theory* (Camerer, Ho and Chong, 2004). This approach assumes that players use a limited number of stages of best-response reasoning, anchoring their beliefs on the assumed behaviour of strategically naïve individuals.

Accumulated experimental evidence suggests that neither approach can give a satisfactory explanation of behaviour across the whole range of coordination games. The main obstacle to a unified explanation seems not to be that *individuals* differ in the modes of reasoning they use. Instead, it is that *games* differ in the modes of reasoning that they evoke. There have been various conjectures about which properties of coordination games facilitate or inhibit the two kinds of reasoning. In particular, it has been suggested that reasoning based on collective optimality may be inhibited if players have conflicting preferences over equilibria, if the collectively optimal solution gives the players unequal payoffs, or if that solution, although Pareto-efficient ex ante, is Pareto-dominated ex post (Crawford, Gneezy and Rottenstreich, 2008; Bardsley et al., 2010). This paper reports experimental tests which support these conjectures. We conclude that behaviour in coordination games is best

explained as resulting from two fundamentally distinct modes of reasoning, one based on collective optimality, the other on bounded best responses; which mode is most likely to be used depends on particular properties of the game being played.

We study a class of two-player *diagonal coordination games*, defined as follows. Each player (P1 or P2) chooses from the same set  $\{s_1, \dots, s_m\}$  of pure strategies, where  $m \geq 2$ . If strategy  $s_k$  is chosen by both players, P1's payoff is  $z_{1,k} > 0$  and P2's payoff is  $z_{2,k} > 0$ ; otherwise, both payoffs are zero. Thus, there are  $m$  pure-strategy equilibria along the main diagonal of the payoff matrix; all off-diagonal payoffs are zero. This class includes *pure coordination games*, defined by  $z_{i,k} = z_{j,l}$  for all  $i, j, k, l$ , and *Hi-Lo games*, defined by  $z_{i,k} = z_{j,k}$  for all  $i, j, k$  and  $z_{i,k} \neq z_{i,l}$  for some  $i, k, l$ . If there are strategies  $j, k$  such that  $z_{1,j} > z_{1,k}$  and  $z_{2,j} < z_{2,k}$ , there is *conflict of interest* between the players with respect to the corresponding equilibria, as in Battle of the Sexes games. Diagonal coordination games present equilibrium selection problems in their purest form, since the only payoff information by which pure-strategy equilibria can be differentiated from one another is contained in the equilibrium payoffs themselves.<sup>1</sup> These games are relatively easy for experimental subjects to understand, particularly if the off-diagonal payoffs are represented as zero increments of money, because the amount of information that players have to process is relatively small, and because, even in games with a large number of strategies, the strategic structure is very simple.<sup>2</sup> We focus on games in which, for each player, any differences between the positive payoffs on the main diagonal are small relative to differences between these payoffs and zero. Thus, players have a strong incentive to coordinate.

We simplify further by using experimental coordination games in which strategies are not labelled in ways that make any equilibrium strongly salient, independently of its payoffs. It is well known that, in pure coordination games in which one equilibrium is uniquely

---

<sup>1</sup> It is possible that, by focusing players' attention on equilibrium payoffs, diagonal coordination games may make team reasoning more salient, and strategic reasoning less salient, than in coordination games with non-zero off-diagonal payoffs. Even so, we think it is a good research strategy to begin with the simplest games in which the two theoretical approaches have divergent implications.

<sup>2</sup> In contrast, the  $3 \times 3$  and  $4 \times 4$  payoff matrices used by Colman, Pulford, and Rose (2008) and Colman, Pulford, and Lawrence (2014) to test cognitive hierarchy and team reasoning theories seem much less easy for subjects to think about. In these experiments, there is some evidence of team reasoning, but a large proportion of subjects behave as level-0 or level-1 cognitive hierarchy reasoners (see Sugden, 2008 for this interpretation of the results in the earlier paper). Since level-0 behaviour is random and level-1 reasoning is effectively non-strategic, these results suggest that subjects' cognitive abilities may have been over-taxed.

salient, players are often successful in coordinating on that equilibrium as a ‘focal point’ (Schelling, 1960; Mehta, Starmer and Sugden, 1994). We minimise the effects of label salience so as to allow a more focused investigation of the effects of variations in payoffs. Nevertheless, such an investigation may throw light on the reasoning by which players use labels to find focal points. Most theoretical accounts of this reasoning follow one or other of the two approaches we have outlined. Some explanations work by transforming a labelled pure coordination game into a Hi-Lo game in which the focal equilibrium payoff-dominates the other equilibria (Bacharach, 1993; Sugden, 1995; Casajus, 2000). Other explanations use a hierarchy of cognitive levels, as in level- $k$  theory, and assume that the non-strategic choices of the least sophisticated players tend to favour focal labels (Lewis, 1969: 24–36; Bacharach and Stahl, 2000; Crawford, Gneezy and Rottenstreich, 2008).

We begin with a description of how the two approaches explain behaviour in coordination games and with a brief review of existing evidence (Section 1). We then present our experimental design. We derive predictions from two theories – specific versions of team-reasoning and level- $k$  theory – which exemplify the collective optimality and bounded best response approaches. Formalising suggestions that have been made in previous literature, we offer three conjectures about how particular properties of diagonal coordination games might facilitate or inhibit reasoning about collective optimality. We derive implications of these conjectures that can be tested in our experiment (Section 2). Our results reveal that, as expected, neither theory is consistently successful, and confirm the implications of two of the three conjectures (Section 3). We show that this conclusion would still stand if other recognised exemplars of the collective optimality and bounded best response approaches were used (Section 4). In the light of these findings, we suggest that it is unlikely that any single theory will be found to explain behaviour in coordination games (Section 5).

## **1. Existing theory and evidence**

### *1.1 Payoff dominance and team reasoning*

Game theorists have often invoked the principle that rational players, faced with a problem of selecting among Nash equilibria, would reject any equilibrium that was strictly Pareto-dominated by another equilibrium. This is the principle of *payoff dominance* (Harsanyi and Selten, 1988: 80–82).

In using this principle, account needs to be taken of isomorphisms in the games to which it is applied. In pure game theory, it is conventional to make the implicit assumption that players lack a common language for identifying strategies, other than by referring to mathematical properties of the payoff matrix. Given the presupposition that game-theoretic recommendations are common knowledge, this assumption implies that strategies that are isomorphic in the sense of Harsanyi and Selten (1988: 73–76) are chosen with equal probability. For example, consider the diagonal coordination game  $G = \langle (10, 10), (10, 10), (9, 9) \rangle$  (that is, a game with  $m = 3$  in which the payoff combinations along the main diagonal are  $(10, 10)$ ,  $(10, 10)$  and  $(9, 9)$ ). Because the two  $(10, 10)$  equilibria are isomorphic with one another, neither of them is *attainable* as the determinate outcome of a recommendation that is common knowledge. A 50:50 probability mix of these two equilibria *is* attainable, but (if payoffs are defined in utility units) this mix is payoff-dominated by  $(9, 9)$  (Crawford and Haller, 1990; Blume, 2000). If in addition there is no common-language labelling of players, a similar analysis applies to isomorphisms between players. For example, in the game  $G' = \langle (10, 9), (9, 10), (9, 9) \rangle$ ,  $(10, 9)$  and  $(9, 10)$  are isomorphic with one another, and so  $(9, 9)$  is payoff-dominant.

The theory of *team reasoning* provides an alternative explanation of why payoff-dominant equilibria are chosen in coordination games. This theory was first proposed as a formalisation of rule utilitarianism by Hodgson (1967) and Regan (1980); it was later developed by Sugden (1993) and Bacharach (1999, 2006). In this paper, we use Bacharach's version of the theory, in which each player chooses her component of the *team-optimal* profile of strategies – that is, the combination of strategies that gives the best payoff to the players collectively (or 'as a team'). The implications of team reasoning and payoff dominance coincide for pure coordination and Hi-Lo games, but can diverge if a coordination game has conflicts of interest.

Bacharach (2006, pp. 87–88, 145) assumes that when (in a game for two players, P1 and P2) each player 'identifies' with the group of both players, each recognises a common *group utility* function  $U(u_1, u_2)$ , where  $u_1$  and  $u_2$  are the utility payoffs to P1 and P2 individually; this function is symmetric (i.e.  $U(u', u'') = U(u'', u')$  for all  $u', u''$ ) and is strictly increasing in both arguments. In applying the theory to experimental games that are played by anonymous players who have common knowledge only of *material* payoffs, we will assume that group utility is symmetric, strictly increasing and weakly concave in material payoffs. The symmetry assumption seems unavoidable, given that the group utility function

is assumed to be common knowledge and the game is anonymous. Weak concavity in material payoffs implies that group preferences are weakly risk-averse (as assumed by Bardsley et al., 2010) and weakly inequality-averse.

Some theorists have proposed models in which players can engage in team reasoning at different levels of cognitive sophistication, analogous with the levels of individual reasoning in level- $k$  theory. In these models, *naïve* team reasoners choose their components of *what they believe to be* the team-optimal strategy profile, but they may fail to recognise certain ways of achieving team optimality. *Sophisticated* team reasoners are not subject to these limitations of rationality, but are aware that their co-players may be. Thus, sophisticated team reasoners choose strategies that are team-optimal when chosen by all players like themselves, given the behaviour of their naïve counterparts (Bacharach, 1993; Bacharach and Bernasconi, 1997; Blume and Gneezy, 2010).

It is generally accepted among theorists of team reasoning that team reasoning is used only when individual players identify with the players as a group, and that group identification occurs only under particular conditions. Conditions that have been hypothesised include the existence of a Pareto-dominated Nash equilibrium (Bacharach, 2006), the absence of individual incentives to deviate from team-optimal solutions (Smerilli, 2012), and positive correlation between players' payoffs (Tan and Zizzo, 2008). Diagonal coordination games necessarily satisfy the first two conditions. In all the games we investigate, there is very strong positive payoff correlation. Thus, these games are particularly suitable for investigating how far players use team reasoning in situations in which group identification might be expected to occur.

## 1.2 *Level- $k$ and cognitive hierarchy theory*

Level- $k$  theory is a theory of non-equilibrium individual reasoning. It assumes that players are of different types, defined by the level of sophistication with which they reason. At the lowest level, 'level 0' or  $L0$ , players' decisions are non-strategic (or, in some applications, strategically 'naïve').<sup>3</sup> Players at each higher level are assumed to know the probability distributions of the decisions of lower-level players, and to choose best responses to the

---

<sup>3</sup> In principle, it would be possible to interpret 'strategic naïveté' to include *any* mode of reasoning that differed from conventional game theory. In this way, principles of team reasoning could be incorporated into the specification of  $L0$  behaviour, and level- $k$  theory might then be said to 'explain' why, in a coordination game, higher-level players replicate this behaviour. We suggest that such a permissive interpretation of naïveté would trivialise level- $k$  theory.

decisions of players at the level immediately below their own. Cognitive hierarchy theory differs by assuming that players at each level above  $L0$  choose best responses to a probability mix of the decisions of all levels below their own.

In level- $k$  modelling, the default assumption is that  $L0$  choices have a uniform random distribution over available pure strategies; the relative frequency of  $L0$  types is often, but not always, assumed to be zero (Crawford, Costa-Gomes and Iriberry, 2013: 14). The former assumption implies that, in a two-player game, an  $L1$  player chooses the strategy with the highest mean payoff to herself (averaging across her opponent's strategies), which might be interpreted as a plausible representation of strategic naïveté. In applying level- $k$  theory to diagonal coordination games with salient labels, Crawford, Gneezy, and Rottenstreich (2008; henceforth 'CGR') assume that  $L0$  players choose 'payoff-salient' strategies, and use 'label salience' to break ties if two or more strategies are payoff-salient. They do not give a general definition of payoff salience but, on the most natural interpretation, an  $L0$  player who acts on payoff salience behaves just like an  $L1$  player who best-responds to uniformly random  $L0$  choices. Thus, when strategies do *not* have salient labels, the level- $k$  model used by CGR reduces to the default model except for the numbering of player types and the absence of any players who choose at random.

### 1.3 Existing evidence

Neither team reasoning nor level- $k$  theory succeeds in organising the existing evidence of behaviour in coordination games.

Proponents of the theory of team reasoning have pointed particularly to two bodies of evidence. The first is that the payoff-dominant equilibrium is almost always chosen in two-strategy Hi-Lo games (e.g. Bardsley et al., 2010). Intuitively, it is hardly surprising that people behave in this way; but this fact requires a theoretical explanation, and the theory of team reasoning provides one. However, the same observation can also be explained by level- $k$  theory.

In an attempt to discriminate between these explanations, Bardsley et al. (2010) ran two experiments using Hi-Lo games with *nondescript* labels in the sense of Bacharach (2006: 16) – that is, labels that are distinct, but do not differ in ways that can be described using

predicates that come to mind easily.<sup>4</sup> The two experiments (run in Amsterdam and Nottingham respectively) had similar designs but used different labels. A typical example of Bardsley et al.'s games has the array of payoff vectors  $\langle (10, 10), (10, 10), (10, 10), (10, 10), (9, 9) \rangle$ . Because of the isomorphism between the  $(10, 10)$  equilibria,  $(9, 9)$  is team-optimal, but the strategies leading to  $(9, 9)$  would not be chosen by level- $k$  reasoners at  $L1$  and above. In the Amsterdam experiment, large majorities of subjects chose team-optimal strategies in games of this type, but in the Nottingham experiment, contrary to both theories, all strategies were chosen with approximately the same frequency. In both experiments, the Hi-Lo games were interspersed with pure coordination games, and the latter used different labels in the two experiments. Bardsley et al. speculate that there was some tendency for the modes of reasoning used in pure coordination games to spill over to Hi-Lo games. Because the focal points in the Amsterdam pure coordination games were labelled in ways that made them 'odd ones out', this may have primed players to think of the unattractive uniqueness of the  $(9, 9)$  equilibrium as a means of coordination. The suggestion is that, in the absence of such priming, players may fail to recognize the focality of an equilibrium that is Pareto-dominated *ex post*. This is one of the conjectures that our experiment was designed to test.

The second body of evidence is of players using salient labelling features to identify focal points in pure coordination games. As we noted in the Introduction, some explanations of the underlying reasoning assume that players use the payoff dominance principle or team reasoning, while others assume level- $k$  reasoning. Much of the evidence from pure coordination games is consistent with both theories, but some evidence suggests that the salient properties that are used to solve coordination problems are not always the same as the properties that stand out for indifferent or naïve individuals, and that might therefore be expected to be favoured by  $L0$  players. The latter properties can be elicited in non-strategic 'just pick' problems, in which choice options are identical except for labelling and lead with certainty to exactly the same outcome (Bardsley et al., 2010).<sup>5</sup> Intuitively, one might expect non-strategic responses to favour labels with positive connotations (for example, 'smiley')

---

<sup>4</sup> Bardsley and Ule (2017) use 'risky coordination games' to discriminate between the two explanations. In these games, as in Hi-Lo games, the payoff matrix is symmetrical, the players' interests are perfectly aligned, and the pure-strategy equilibria (along the main diagonal) are Pareto-ranked. However, the off-diagonal payoffs are such that level- $k$  reasoners would not reach the team-optimal solution. Bardsley and Ule's experimental results support the theory of team reasoning.

<sup>5</sup> Bardsley et al. interpret this as evidence in support of team reasoning and contrary to level- $k$  theory. For an alternative interpretation, see Crawford, Costa-Gomes, and Iriberry (2013: 48–49).



faces rather than ‘sad’ ones), but labels with uniquely negative connotations are often effective as focal points (Rubinstein, Tversky and Heller, 1996). When odd-one-out labels are focal points in pure coordination games,  $L0$  specifications which assume a bias towards such labels do not perform well when applied to other types of game (Hargreaves Heap, Rojo Arjona and Sugden, 2014). Taken together, the evidence points to some form of reasoning in which each player looks for a solution that is good ‘for us’, and which recognises uniqueness among labels as an effective coordination device. In other words, something like team (or payoff-dominance) reasoning seems to be implicated in players’ use of focal points in pure coordination games.

However, there is also evidence suggesting that the power of focal points is restricted to pure coordination games. CGR report experiments on ‘X-Y’ games – diagonal coordination games with  $m = 2$ , where the equilibria are isomorphic to one another but one has a uniquely salient label. They find that players coordinate with a high degree of success when these are pure coordination games, but that when there is even a small degree of conflict of interest, coordination rates are no better than would result from random choice. Isoni et al. (2013) investigate a class of ‘tacit bargaining’ games that are formally equivalent to X-Y games but framed differently. In these games, coordination is significantly better than random, even when there are conflicts of interest (contrary to CGR’s level- $k$  model), but conflicts of interest still reduce coordination rates (contrary to the theory of team reasoning). The tendency for conflicts of interest to weaken the power of focal points is inconsistent with team reasoning, even if one allows for the possibility that some team-reasoning players are naïve. (In order to use a focal point in a pure coordination game, a team-reasoning player must be able to recognise the uniqueness of its label, and to recognise that this uniqueness can be used as a means of coordination. A player with these capacities would presumably be capable of using the same focal point in the presence of a small conflict of interest.) However, this pattern can be explained by a level- $k$  model in which  $L0$  choices are governed by payoff and label salience.

CGR also report experiments on a class of ‘pie games’. These are diagonal coordination games with  $m = 3$ , where one strategy has a uniquely salient label. (Equilibria are represented as slices in a pie diagram; two slices have the same colour, and the third slice is coloured differently.) One of these games, ‘AM1’, can be described by the array of payoff vectors  $\langle (6, 5), (5, 6), (5, 5)^* \rangle$ , where the asterisk denotes the salient label. Game ‘AL1’ is described by  $\langle (10, 5), (5, 10), (5, 5)^* \rangle$ . In AM1 and AL1, the strategy with the salient label

was chosen by 90 per cent and 83 per cent of subjects respectively, consistently with team reasoning but not with CGR's level- $k$  model.<sup>6</sup> However, the theory of team reasoning failed in game 'AM3', described by  $\langle(6, 5), (5, 6), (7, 5)^*\rangle$  and in game 'AM4', described by  $\langle(6, 7), (7, 6), (7, 5)^*\rangle$ . In both these games, team reasoning theory predicts the choice of the saliently labelled strategy, but this choice was made by only 29 per cent of subjects in AM3 and by only 7 per cent in AM4.<sup>7</sup>

Notice that this body of evidence cannot be explained simply by hypothesising that players try to coordinate on focal points, and that the determinants of focality differ across games. If the concept of focality (in a two-player game) is to be used in a non-trivial way, it must refer to some property by virtue of which *a pair of* strategies stands out for both players, and which each of them uses *with the intention of coordinating with the other*.<sup>8</sup> In most of the cases in which players of coordination games behave consistently with level- $k$  theory but not with team reasoning, that behaviour is not plausibly explained as an attempt to reach a focal point. For example, consider CGR's X-Y games, played by University of Chicago students, in which the pure-strategy equilibria are labelled 'Sears Tower' and 'AT&T Building'. It seems undeniable that the 'Sears Tower' label, which refers to a prominent Chicago landmark, stands out for both players, irrespective of whether the equilibrium payoffs for the players are equal or unequal. The effect of introducing small payoff inequalities is not to make a different equilibrium stand out as a means of coordination; it is to make players less likely to *use* the fact that 'Sears Tower' stands out when reasoning about which strategy to choose, and hence to fail to coordinate at all.

The evidence we have reviewed suggests that two distinct modes of reasoning – one based on bounded best responses, the other on collective optimality – are used in coordination games. A plausible conjecture, compatible with the evidence from AM3 and

---

<sup>6</sup> Bett et al. (2016) report similar results from experiments on diagonal coordination games that are similar to AL1 and AL2 except for the absence of salient labels. In games described by  $\langle(6, 7), (7, 6), (x, x)\rangle$ , the proportion of subjects choosing the strategy consistent with  $(x, x)$  when  $x$  equalled 6.5, 6 and 5 was respectively 100 per cent, 100 per cent, and 77.8 per cent.

<sup>7</sup> The relevant predictions of team reasoning theory are explained in Section 2.2 below.

<sup>8</sup> Game theorists sometimes use the term 'focal point' to describe *any* equilibrium (in a game with multiple equilibria) on which players in fact succeed in coordinating. Clearly, this *ex post* concept of focality has no explanatory value. Another usage, specific to level- $k$  and cognitive hierarchy theory, is to define an equilibrium as a focal point if the corresponding strategies are chosen with high probability by  $L0$  players. But the evidence we have cited shows that, in some coordination games, there is successful coordination on strategies that are *not* favoured at  $L0$ .

AM4 and from X-Y games with conflicts of interest, is that reasoning about collective optimality is inhibited if the players' material payoffs in the team-optimal equilibrium are unequal. CGR (p. 1456) invoke a similar idea when they suggest that the use of team reasoning is negatively affected by 'payoff conflict'. Our experiment was designed to test this conjecture.

## 2. Experimental design and hypotheses to be tested

### 2.1 *The games used in the experiment*

In our experiment, each subject played eighteen diagonal coordination games, modelled on CGR's pie games. Each game is defined by an array of three strategies  $\langle s_1, s_2, s_3 \rangle$  for each of players P1 and P2. In each game, the corresponding array of payoff vectors is  $\langle (x, y), (y, x), (v, w) \rangle$ , satisfying  $x \geq y > 0$ ,  $v \geq w > 0$  and  $\{x, y\} \neq \{v, w\}$ . This definition imposes the restriction that the payoff vectors for the  $s_1$  and  $s_2$  equilibria are symmetric with one another, but that both are distinct from that for the  $s_3$  equilibrium.<sup>9</sup>

Figure 1 shows an example of the *payoff display* that was used to describe games to subjects. The game used in this example (G2 in the experiment) has  $x = 10$ ,  $y = 9$ ,  $v = 8$  and  $w = 7$ ; payoffs are in euros. (The strategy labels ' $s_1$ ', ' $s_2$ ' and ' $s_3$ ' were not seen by subjects, but have been added to aid interpretation.) The payoff display was made up of three circles, which we will call 'left', 'right' and 'bottom'. The text in each circle showed the payoffs in one of the three pure-strategy equilibria, referring to the players only as 'you' and 'the other' to ensure that there was no commonly-known and payoff-independent labelling by which the players could be separately identified. All circles were the same colour (a feature that differentiates our experiment from CGR's). There were three experimental treatments, A, B, and C, which differed only in respect of which equilibrium was shown in which position. In treatment A, the payoffs of the  $s_1$ ,  $s_2$  and  $s_3$  equilibria were shown to both players in the left, right and bottom circles respectively. The payoff displays for treatments B and C were generated by rotating the treatment A display counterclockwise through 120 and 240 degrees. In each game, each player independently chose one of the circles by clicking on it. If they

---

<sup>9</sup> Six of our games have  $x > y$  and  $v > w$ . In these games,  $s_1$  and  $s_2$  are not completely isomorphic with one another, because the player who gets the higher payoff in the  $s_1$  (respectively:  $s_2$ ) equilibrium gets the higher (respectively: lower) payoff in the  $s_3$  equilibrium. Such asymmetries can be relevant for level- $k$  reasoning, but we thought it unlikely that they would be salient for players who were looking for a focal point for coordination. We discuss this issue further in Section 3.1.

both chose the same circle, their payoffs were as stated on that circle; otherwise the payoff to both players was zero.

[Figure 1 near here]

Notice that the only labelling feature that distinguished  $(x, y)$  from  $(y, x)$  was the positioning of the corresponding circles in the payoff display. Our working assumption, which we will call *positional non-salience*, was that this feature would not be sufficiently salient for players to be able to use it as an effective coordination device. However, by using the three different treatments, we were able to investigate whether strategy choices were influenced by the positioning of the circles.

Some readers may be surprised that we did not use *scrambled* labels (in the sense of Crawford and Haller, 1990), for example, by randomising the three rotations of the payoff display independently for each player in each game. But in such a design,  $(x, y)$  and  $(y, x)$  are no longer the payoff vectors of two equilibria that result from distinct (although isomorphic) combinations of strategy choices by the two players. Each player has only two strategically meaningful options – *either* to choose the  $(v, w)$  circle, *or* to choose one of the other two. If both players make the first choice, the outcome is  $(v, w)$ ; if both make the second choice, and irrespective of which of the two relevant circles each player clicks on, the outcome is a 0.5:0.5 probability mix of  $(x, y)$  and  $(y, x)$ . For the values of  $v, w, x$  and  $y$  used in our experiment, the effect of scrambling would be to reduce our 3×3 games to 2×2 Hi-Lo games. Since team reasoning and level- $k$  theories make the same predictions for Hi-Lo games, that would defeat the object of our experiment.

The payoff vectors of the games G1–G18 used in the experiment are shown in Table 1. The rows and columns of this table provide a structure for classifying the games. This structure reflects the conjectures that the experiment was designed to test. Each row contains games characterised by a different combination of equality ( $x = y, v = w$ ) or inequality ( $x > y, v > w$ ) in the equilibrium payoffs. Each column contains games characterized by a different relation between the equilibrium payoffs. The  $s_3$  equilibrium is *strictly Pareto-dominated* by both the  $s_1$  and  $s_2$  equilibria in games with  $y > v$ . It is *weakly Pareto-dominated* if  $y = v$  and either  $x > y$  or  $v > w$ ; it is *weakly Pareto-dominating* if  $w = x$  and either  $x > y$  or  $v > w$ ; and it is *strictly Pareto-dominating* if  $w > x$ . If either  $v > x \geq y > w$  or  $x > v \geq w > y$ , the  $s_3$

equilibrium is *not Pareto-comparable* with the other equilibria.<sup>10</sup> Because of the requirement that the  $s_3$  equilibrium has a distinct payoff vector, the combination of  $x = y$  and  $v = w$  implies that the  $s_3$  equilibrium is either strictly Pareto-dominated or strictly Pareto-dominating, and so three cells of Table 1 are necessarily empty. If  $x > y$  and  $v > w$ , Pareto non-comparability implies either  $x > v > w > y$  (making the  $s_3$  equilibrium unambiguously more equal than the others) or  $v > x > y > w$  (with the opposite effect). As we were interested in whether considerations of payoff equality affected subjects' propensities to use team reasoning, we included games representing both these cases (G9 and G10). Apart from this exception, there is exactly one game in every cell that is consistent with the general restrictions we have imposed on payoffs.

[Table 1 near here]

Since our objective was to investigate the respective conditions under which individuals are most likely to use collective-optimality and bounded best-response modes of reasoning, we need to know what each of these modes imply about behaviour in our games. Each of them can be formulated in a range of alternative theoretical models. We will focus on two specific theories, representative of the two approaches. After reporting our findings in relation to these theories, we will consider whether our main conclusions are sensitive to our choice of representative theories.

In the following two sections, we present predictions derived from the representative theories. We emphasise that our prior expectation, based on the existing evidence summarised in Section 1.3, was that neither theory would perform well across all eighteen games. Our interest was in the *pattern* of successes and failures that we would find.

## 2.2. Predictions of the theory of team reasoning

We use Bacharach's (1999, 2006) theory of team reasoning as the representative of the collective optimality approach. We assume that the group utility function is symmetrical, increasing and weakly concave in material payoffs (see Section 1.1 above).

The payoff vectors of games G1–G18 were chosen so that, given that there is positional non-salience, the choice of  $s_3$  by both players is the uniquely team-optimal strategy.

---

<sup>10</sup> There are two further possibilities, not represented in our experiment: (i) that the  $s_3$  equilibrium Pareto-dominates one but not both of the other equilibria, and (ii) that the  $s_3$  equilibrium is Pareto-dominated by one but not both of the other equilibria.

To see why, notice that if neither  $s_1$  nor  $s_2$  has a uniquely salient distinguishing feature, the team-optimal choice must be *either* (i) that both players choose  $s_3$  or (ii) that each player independently randomises between  $s_1$  and  $s_2$ , choosing each of these strategies with probability 0.5. Let  $Z(z_1, z_2)$  be the group utility function defined with respect to the players' material payoffs  $z_1$  and  $z_2$ , normalized so that  $Z(0, 0) = 0$ . Then group utility is  $Z(v, w)$  in case (i) and (using the symmetry assumption)  $Z(x, y)/2$  in case (ii). Given that  $x \geq y > 0$  and  $v \geq w > 0$ , it is sufficient for unique team-optimality of (i) that  $Z(w, w) > Z(x, x)/2$ . Since, by weak convexity,  $Z(x/2, x/2) \geq Z(x, x)/2$ , it is sufficient that  $Z(w, w) > Z(x/2, x/2)$  or (using increasingness)  $w > x/2$ . In our games, the value of  $w$  is never less than  $0.7x$ .<sup>11</sup> Thus, the theory of team reasoning (as we have formulated it) predicts that, in every game, the proportion of  $s_3$  choices is 1 for both players. These implications are recorded in the first column of Table 2.

[Table 2 near here]

In games G9 and G11–G18, the assumed properties of group utility imply  $Z(v, w) \geq Z(x, y) = Z(y, x)$ . This inequality is strict for G12–G18, and is also strict for G9 and G11 if group utility is *strictly* concave. Thus, since the  $s_3$  equilibrium can be uniquely identified by its payoffs alone,  $s_3$  is team-optimal in these games independently of how strategies are labelled. In the remaining games, the conclusion that  $s_3$  is team-optimal could in principle be reversed if either  $s_1$  or  $s_2$  had a uniquely salient label. However, for  $s_1$  or  $s_2$  to be team-optimal, the salience of its label must be extremely pronounced. For example, consider G3 =  $\langle (10, 10), (10, 10), (9, 9) \rangle$ . Suppose that if both players follow the rule 'Choose whichever of  $s_1$  or  $s_2$  has the more salient label', each player identifies  $s_1$  as 'more salient' with independent probability  $q$ , where  $q > 1/2$ . Then the probability that this rule will lead to coordination is given by  $q^* = q^2 + (1 - q)^2$ . Given our assumptions about group utility,  $s_1$  is team-optimal if and only if  $q^* \geq 0.9$ , or equivalently,  $q \geq 0.95$ . Since it seemed unrealistic to expect this degree of correlation between players' judgements about which of two positions in the payoff display was more salient, we thought it reasonable to assume positional non-salience when designing tests of the theory of team reasoning. However, if significant number of players do *not* choose  $s_3$ , our design allows us to investigate whether (for subjects as a whole, and/or at

---

<sup>11</sup> In our notation, CGR's games AM1, AM3 and AM4 satisfy the condition  $w > x/2$ , making  $s_3$  team-optimal independently of the salience of its label. In their AL1 game, the  $s_3$  equilibrium  $(5, 5)^*$  is team-optimal if  $Z(5, 5) > Z(10, 5)/2$ . This inequality is implied by symmetry, increasingness and weak concavity.

the individual level) there are systematic asymmetries in the distribution of choices between  $s_1$  and  $s_2$ .

In testing the theory of team reasoning, our null hypothesis is that subjects choose strategies at random. Our alternative hypothesis is thus:

*Hypothesis 1 (team reasoning):* In each of games G1–G18, the aggregate proportion of  $s_3$  choices is greater than  $1/3$ .

This hypothesis is conservative: one might reasonably argue that if the theory of team reasoning were ‘really’ true, the aggregate proportion of  $s_3$  choices in each game would be much greater than  $1/3$ . No deterministic theory of human decision-making can be expected to yield absolutely correct predictions, but one might at least hope that such a theory would out-predict the hypothesis that individuals choose at random.

### 2.3 Predictions of level- $k$ theory

Our representative bounded best-response theory is level- $k$  theory, specified such that  $L0$  choices are uniformly distributed over the three strategies. The only assumption we make about the distribution of levels of reasoning in the population is that the proportion of players at  $L1$  or higher is greater than 0. We also assume that utility payoffs are the same as (or linear transformations of) material payoffs. The proportions of  $s_3$  choices predicted by this specification of level- $k$  theory, disaggregated by players’ levels of reasoning, are shown in Table 2. The derivation of these predictions is explained in Appendix 1 [intended for online publication].<sup>12</sup>

As in our tests of team reasoning, our null hypothesis is that subjects choose strategies at random:

*Hypothesis 2 (level- $k$  reasoning):* In each of games G1–G7, G9 and G11, the aggregate proportion of  $s_3$  choices is less than  $1/3$ . In each of games G8, G10 and G12 – G18, the aggregate proportion of  $s_3$  choices is greater than  $1/3$ .

This hypothesis, like Hypothesis 1, is conservative. Nevertheless, the two hypotheses contradict one another for nine of the games we investigate. This leaves plenty of room for each of the theories to fail.

---

<sup>12</sup> Level- $k$  theory has additional implications, shown in Appendix 1 but not in Table 2, about the distribution of choices between  $s_1$  and  $s_2$ . Since the theory of team reasoning has no corresponding implications (it predicts that  $s_1$  and  $s_2$  are not chosen), we focus on predictions about  $s_3$  choices.

## 2.4 Three conjectures

Our objective was to test conjectures about the conditions under which each theory was more or less likely to work well. We will formulate these conjectures in terms of conditions that might facilitate or inhibit team reasoning – that is, make team reasoning more or less likely. But this is merely a matter of wording: in the context of our investigation, facilitating team reasoning is the same thing as inhibiting level- $k$  reasoning (and conversely). Our background assumption, supported by the existing evidence reviewed in Section 1.3, is that behaviour in diagonal coordination games is explained by some (as yet, not fully known) combination of the two theories.

The payoff vectors of the games in the experiment were chosen to allow us to investigate the following three conjectures, informed by the previous literature:

*Conjecture 1:* Other things being equal, equality of material payoffs in the team-suboptimal equilibria (i.e.  $x = y$ ) inhibits team reasoning.

In terms of Table 1, Conjecture 1 proposes that within any column (that is, holding constant the Pareto-dominance relationships between equilibria), team reasoning is more likely in a game in the second row than in a game in the first, and more likely in a game in the fourth row than in a game in the third. The underlying intuition is that it is psychologically easier to think of oneself as playing one's part in a joint action if the alternatives to so acting lead to ex post payoff inequality.

*Conjecture 2:* Other things being equal, equality of material payoffs in the team-optimal equilibrium (i.e.  $v = w$ ) facilitates team reasoning.

Thus, within any column of Table 1, team reasoning is more likely in a game in the third row than in a game in the first, and more likely in a game in the fourth row than in a game in the second. The underlying intuition is that it is psychologically easier to think of oneself as playing one's part in a joint action if all participants benefit from it equally. Given that the payoff vectors of the  $s_1$  and  $s_2$  equilibria are symmetric, as in our set-up, it is natural to think of  $v = w$  as equivalent to 'the players benefit equally from their both choosing  $s_3$ '. This equivalence would be less obvious in a diagonal coordination game such as  $\langle (10, 5), (10, 5), (11, 11) \rangle$ , in which the team-suboptimal equilibria are not symmetric.

*Conjecture 3:* Other things being equal, team reasoning is inhibited if, ex post, each of the team-suboptimal equilibria strictly or weakly Pareto-dominates the team-optimal equilibrium (i.e.  $[y > v]$  or  $[x > y = v]$  or  $[y = v > w]$ ).



Thus, within any row of Table 1, team reasoning is more likely in a game in the third column than in a game in the first or second. The intuition is that, when two equilibrium payoff vectors are isomorphic and neither has a uniquely salient label, some individuals may lack the cognitive sophistication to recognise that the best attainable strategy for the players jointly is to choose the uniquely *unattractive* equilibrium.

Conjectures 1–3 are formulated in terms of the modes of reasoning that players use, not the choices that they ultimately make. Since our experimental design does not allow us to elicit modes of reasoning independently of decisions, these conjectures can be tested only in terms of their implications for players’ strategy choices.<sup>13</sup> More specifically, these implications refer to variations in the proportion of  $s_3$  choices between pairs of games which differ along particular dimensions.

Consider any pair of games ( $G, G'$ ) such that (‘condition 1’) exactly one of Conjectures 1–3 proposes that team reasoning is more likely in  $G'$  than in  $G$  (and, by implication, that level- $k$  reasoning is less likely). Assume (‘condition 2’) that level- $k$  theory predicts that  $s_3$  choices are equally or less likely in  $G'$  than in  $G$ . Then if the relative frequencies of the two modes of reasoning were constant across the two games,  $s_3$  choices would be equally or less likely in  $G'$ . (Recall that the theory of team reasoning predicts that  $s_3$  is chosen with probability 1 in all games.) Assume also (‘condition 3’) that level- $k$  theory implies that, in  $G$ ,  $s_3$  is chosen with probability less than 1 by players at some level higher than  $L0$ . Then, if the relevant conjecture were true, the greater use of team reasoning in  $G'$  might induce a higher proportion of  $s_3$  choices than in  $G$ . Thus, pairs of games that satisfy all three conditions can be used to test whether the proportion of  $s_3$  choices changes from  $G$  to  $G'$  in a direction that is consistent with a specific conjecture but not consistent with the assumption that each mode of reasoning is used with constant frequency. Using all such pairs of games, we construct the following hypotheses:

*Hypothesis 3.1 (consistency with Conjecture 1):* In each of the pairs of games (G1, G2), (G3, G4), (G5, G6), (G8, G9) and (G8, G11), the proportion of  $s_3$  choices is higher in the second game than in the first.

---

<sup>13</sup> It is possible to use experimental designs which elicit modes of reasoning independently of decisions. For example, van Elten and Penczynski (2015) reconstruct players’ modes of reasoning in coordination games from the texts of ‘recommendation’ messages that they send to other players. One of their findings is that ‘TR3’ messages that refer to the difficulty of coordinating on one of two isomorphic equilibria are very uncommon. This finding gives some support to Conjecture 3.

*Hypothesis 3.2 (consistency with Conjecture 2):* In each of the pairs of games (G1, G3), (G2, G4), (G6, G67), (G9, G11), (G10, G11) and (G13, G14), the proportion of  $s_3$  choices is higher in the second game than in the first.

*Hypothesis 3.3 (consistency with Conjecture 3):* In each of the pairs of games (G2, G9), (G6, G9), (G4, G11) and (G7, G11), the proportion of  $s_3$  choices is higher in the second game than in the first.

## 2.5 *Experimental procedures*

A total of 126 subjects participated in the experiment at the CEEL Lab of the University of Trento in April 2014. Seven sessions were conducted, each with eighteen participants. Subjects were undergraduate students (52.0 per cent from economics and management, 50.0 per cent females, 91.6 per cent Italians). The experiment was programmed using the z-Tree platform (Fischbacher, 2007). The instructions for the experiment (reproduced in the original Italian and in English translation in Appendix 2 [intended for online publication]) were provided to subjects in written form, and were also read aloud by the experimenter to ensure that they were common knowledge. Participants' understanding of the instructions was checked using a short questionnaire; we did not proceed with the experiment until all participants had answered all questions correctly.

Each subject played games G1 to G18 in the eighteen consecutive 'rounds' of the experiment. There was no feedback until all eighteen rounds had been conducted. In each round, subjects were matched anonymously; each matched pair of co-players then played one of the games, with the same treatment (A, B, or C, corresponding with different rotations of the payoff display) for both players. One co-player in each pair played as P1, the other as P2; but (as explained in Section 2.1), games were described to subjects only in terms of 'you' and 'the other'. Co-players were rematched between games in such a way that each pair of subjects were co-players in no more than two games. The order in which the games were played by different subjects were counterbalanced to control for order effects. Each subject played nine games as P1 and nine as P2, and played six games in each of the three treatments.<sup>14</sup>

---

<sup>14</sup> The matching of co-players, assignment of games to rounds and treatments, and assignment of subjects to player roles were predetermined for each session. This protocol is explained in Appendix 3 [intended for online publication].

At the end of the experiment, one of the eighteen rounds was randomly selected and subjects were paid according to the outcome of the game they had played in that round. Average earnings for each participant were €6.23. Each subject also received a show-up fee of €3. The average length of a session was approximately 40 minutes.<sup>15</sup>

### 3. Results

In this Section, we consider those features of the results that are most relevant for tests of the two theories and of our three conjectures. The full results of the experiment, with strategy choices disaggregated by game, player role and treatment, are presented in Appendix 4 [intended for online publication].

#### 3.1. Tests for effects of labelling and non-isomorphism

Hypothesis 1, which states implications of the theory of team reasoning, was derived under the assumption that any asymmetries between  $s_1$  and  $s_2$  were insufficiently salient to be effective coordinating devices. We need to check whether our observations are consistent with that assumption.

We do this by investigating how far the aggregate choices of players who did not choose  $s_3$  were skewed towards one or other of  $s_1$  or  $s_2$ . If, contrary to our assumption, it was team-optimal for players to use asymmetries between  $s_1$  and  $s_2$  as a means of coordination, one would expect a very high degree of skew in these choices. The distributions of choices between  $s_1$  and  $s_2$  in each of the eighteen games are shown in Table 3. In games in which  $s_1$  and  $s_2$  are isomorphic (that is, either  $v = w$  or  $x = y$ ), any systematic asymmetries between the frequencies with which they are chosen can be explained *only* as the result of labelling. When  $s_1$  and  $s_2$  are not isomorphic, it is also conceivable that players could coordinate on one of them by using payoff information (for example, by following the rule ‘Choose whichever of  $s_1$  or  $s_2$  favours the player who is also favoured by  $s_3$ ’). For each game and each treatment, we test the null hypothesis that each of  $s_1$  and  $s_2$  is equally likely to be chosen; we report the  $p$ -value for a two-tail binomial test. For games in which the two strategies are not isomorphic, we also report this test for data pooled across all three treatments, to check

---

<sup>15</sup> The experiment reported in this paper was preceded by a pilot experiment, using a smaller set of games and with different subjects. The pilot is reported by Faillo, Smerilli and Sugden as ‘The roles of level- $k$  and team reasoning in solving coordination games’, Working Paper 6-13 of Cognitive and Experimental Economics Laboratory, University of Trento (2013).

whether players used *payoff* asymmetries between  $s_1$  and  $s_2$  as coordination devices.<sup>16</sup> The pooled tests show no significant asymmetries.

[Table 3 near here]

Although there are relatively few individual game/treatment pairs in which label-based asymmetries are statistically significant, the aggregate data suggest small but systematic asymmetries in treatments A and B. In treatment A,  $s_1$  was shown in the left circle and  $s_2$  in the right. Summing over all eighteen games, there are 347 (non-independent) observations of choices of  $s_1$  or  $s_2$ ; 218 of these (62.8 per cent) are of  $s_1$ . In treatment B,  $s_1$  was shown in the bottom circle and  $s_2$  in the left. Of the 372 choices of  $s_1$  or  $s_2$ , 150 (40.3 per cent) are of  $s_1$ . In contrast, there was no consistent asymmetry in treatment C, in which  $s_1$  was shown in the right circle and  $s_2$  in the bottom, where the 361 choices of  $s_1$  or  $s_2$  were almost equally divided between  $s_1$  (172 choices) and  $s_2$  (189 choices). However, the asymmetries observed in treatments A and B fell far short of the levels that would make deviations from  $s_3$  team-optimal. (For example, the 218: 129 asymmetry in treatment A implies a probability of coordination of only 0.533.)

This leaves the possibility that some subjects had the mistaken belief that, with high probability, their own sense of which of the three positions (left, right or bottom) was most salient would be shared by their co-players. Such subjects might choose  $s_1$  or  $s_2$  in the belief that this was team-optimal, even though in fact it was not. But if this were the case, one would expect that team-reasoning subjects would systematically favour whichever position they thought most salient. We tested, for each subject separately, whether subjects' choices in the eighteen games were distributed non-randomly over the three positions. The null hypothesis of a random distribution was rejected for only nine of the 126 subjects (at 5 per cent significance in a two-sided chi-squared goodness of fit test). It seems that few if any subjects considered the use of position as a means of coordination.

### 3.2 Tests of Hypotheses 1 and 2

The data relevant for tests of Hypotheses 1 and 2 are reported in Table 4.

[Table 4 near here]

---

<sup>16</sup> There are six such games (G2, G6, G9, G10, G13 and G16). For all these games except G13, level- $k$  theory predicts that, when aggregated across players,  $s_1$  and  $s_2$  choices are equally probable, irrespective of the distribution of levels of reasoning (see Appendix 1, Table A1.1).

A striking feature of these data is the wide variation in the proportion of  $s_3$  choices across games, ranging from 1.6 per cent in G2 to 96.8 per cent in G18. In every game, this proportion is significantly different from the random-choice benchmark (in all but two cases, with  $p < 0.01$ ).<sup>17</sup> These differences are as predicted by Hypothesis 1 in eleven games (G4, G7, G9 and G11–G18), but are in the unpredicted direction in the other seven games. They are as predicted by Hypothesis 2 in twelve games (G1, G2, G3, G5, G6 and G12–G18), but are in the unpredicted direction in the other six games. There are two games (G8 and G10) for which both hypotheses are disconfirmed. Thus, in line with the mixed results of previous studies and with our prior expectation, the pattern of observed deviations from the random-choice benchmark cannot be adequately explained by either of the theories we are considering.

### 3.3 Tests of Hypotheses 3.1–3.3

Table 5 lists the fifteen pairs of games that are relevant for tests of Hypotheses 3.1–3.3. For each pair, the table shows which hypothesis implies that the proportion of  $s_3$  choices increases as one moves from the first game in the pair to the second. It also shows the level- $k$  prediction of the change in the proportion of  $s_3$  choices (derived from Table 2). The corresponding team-reasoning prediction is always ‘no change’. In cases in which the level- $k$  prediction is ‘no change’, the hypothesis that is being tested is a strict implication of the corresponding conjecture (since in these cases, an increase in the relative frequency of team reasoning would imply an increase in the proportion of  $s_3$  choices). When the level- $k$  prediction is ‘decrease’, the hypothesis that is being tested is stronger than the corresponding conjecture (since an increase in the relative frequency of team reasoning might be offset by the reduction in  $s_3$  choices by subjects who consistently use level- $k$  reasoning); thus, confirmation of the hypothesis gives particularly strong support to the conjecture. Of course, these fifteen tests are not independent of one another, but each test, considered in isolation, is a valid test of the relevant hypothesis.

*[Table 5 near here]*

---

<sup>17</sup> Our tests are two-tail tests of the null hypothesis that the probability that a randomly-selected player of the relevant game chooses  $s_3$  is  $1/3$ . One might reasonably argue that, for any given game, a game theorist who specifically proposes one of the two theories is entitled to use a one-tail test of the hypothesis that this probability differs from  $1/3$  in the direction predicted by the proposed theory. However, from the viewpoint of an observer who is not proposing any particular theory and simply looking for possible regularities, randomness is the natural null hypothesis, and two-tail tests are appropriate. In fact, the results of our two-tail tests are so strong that one-tail tests are unnecessary.

Conjectures 2 and 3 are very strongly supported by the data. In each of the ten tests of Hypotheses 3.2 and 3.3, there is an increase in the proportion of  $s_3$  choices, consistent with the relevant conjecture. In nine of these cases, the increase is statistically significant, always at the 1 per cent level; these include two cases in which level- $k$  theory predicts a decrease. In contrast, we do not find consistent support for Conjecture 1. In the five tests of Hypothesis 3.1, the conjectured increase in the proportion of  $s_3$  choices is observed in three cases, in two of which it is significant at the 1 per cent level, but there are significant *decreases* in the two other cases.

A comparison between behaviour in G9 and G10 may give further insight into the relationship between payoff inequality and team reasoning. In each of these games, the  $s_1$ ,  $s_2$  and  $s_3$  equilibria all give unequal payoffs and are not Pareto-comparable. The difference between the two games concerns the *degree* of inequality. In G9, the  $s_3$  equilibrium is less unequal than the others; in G10, it is more unequal. Level- $k$  theory implies that  $s_3$  is more likely to be chosen in G10 (see Table 2). However, we observe a much higher proportion of  $s_3$  choices in G9 (69.0 per cent compared with 19.8 per cent, (McNemar's  $\chi^2=45.76$ ,  $p<0.001$ ). This result suggests that team reasoning may be facilitated if the team-optimal equilibrium, although itself unequal, is less unequal than the other equilibria.

#### 3.4. Regression analysis

As a further attempt to identify patterns in the data, we estimate a random effect probit model with a binary dependent variable assuming value 1 if subject  $i$  chooses  $s_3$  in round  $t$  and 0 otherwise. Three regressors capture characteristics of the games which, according to Conjectures 1–3, can affect the probability of choosing  $s_3$ : equality of payoffs in the team-suboptimal equilibria (variable  $x$  equal to  $y$ ), equality of payoffs in the team-optimal equilibrium (variable  $v$  equal to  $w$ ), and having a team-optimal equilibrium that is Pareto-dominated, ex post, by each of the two team-suboptimal equilibria (variable *Pareto dominated*). Additional regressors capture other potentially relevant characteristics of games: having a team-optimal equilibrium that ex post Pareto-dominates each of the team-suboptimal equilibria (variable *Pareto dominating*), playing the game as player P1 (variable *Player P1*), playing a game in which  $s_3$  is either in the left circle (variable  $s_3$  left) or in the right circle (variable  $s_3$  right), and the ratio between the average payoffs of the two team-suboptimal equilibria and the payoffs of the team-optimal equilibrium (variable *Average payoff ratio*). We control also for gender, age and nationality of subjects, the number of

previous experiments in which they have already participated, and the university course in which they are enrolled.

The results are reported in Table 6.

[Table 6 near here]

This analysis provides support to each of the three conjectures. Consistently with Conjectures 1 and 2, it is confirmed that, *ceteris paribus*, the probability of choosing  $s_3$  is positively affected by equality of payoffs in the team optimal equilibrium (variable  $v$  equal to  $w$ ) and negatively affected by equality of payoffs in the suboptimal equilibria (variable  $x$  equal to  $y$ ). Consistently with Conjecture 3, this probability is lower in games in which the  $s_3$  equilibrium is Pareto-dominated ex post by the other two equilibria than in games in which the equilibria are not Pareto-comparable (variable *Pareto dominated*).

Additionally, the probability of choosing  $s_3$  is greater in games in which the  $s_3$  equilibrium Pareto-dominates the other two equilibria ex post than in games in which the equilibria are not Pareto-comparable (variable *Pareto dominated*); and *Average payoff ratio* has a strong and significant negative effect on the probability of choosing  $s_3$ . These effects might be evidence of further factors that affect players' propensity to use team reasoning. It would be plausible to conjecture that team reasoning is more likely when the team-optimal equilibrium is Pareto-dominant, and is also more likely, the higher the average payoff in the team-optimal equilibrium relative to that in the team-suboptimal equilibria. However, the observed effect of Pareto-dominance is also an implication of level- $k$  theory, and *Average payoff ratio* may simply be picking up a general tendency for players to favour strategies with high potential payoffs to themselves individually. We also find a marginally significant negative effect of  $s_3$  being in the right circle. The *Player PI* variable has no significant effect, despite the fact that  $v > w$  in eleven of the eighteen games. Thus, although payoff inequality in the team-optimal equilibrium inhibits team reasoning, there is no evidence that the strength of this effect differs between the two players.

### 3.5. Individual heterogeneity

Recall that there are nine games (G1–G7, G9 and G11) in which level- $k$  and team reasoning have opposite implications about the choice of  $s_3$ .<sup>18</sup> Across these games, there is extreme variation in proportion of  $s_3$  choices, ranging from 1.6 per cent in G2 to 88.9 per cent in G11. Clearly, the vast majority of subjects behaved in accordance with level- $k$  theory in some of these games and in accordance with team reasoning in others. In some of these games, however, there was considerable heterogeneity in individual behaviour. It is therefore relevant to ask whether there is an individual-specific component to players' propensities to use team reasoning and if so, how important that component is.

Given the patterns that we have found in our aggregate data, it would be surprising if there were *no* component of this kind. We have identified certain general properties of games which tend to facilitate or inhibit team reasoning. It would be natural to expect individuals to differ in their susceptibility to these factors. Since our subjects were not selected to be a representative sample of any population, measures of the distribution of 'susceptibility' in the subject pool may not tell us much that is useful about the world outside the lab. However, were we to find that subjects could be divided into discrete types, each with a sharply distinct mode of behaviour, that might provide useful clues for future research.

As a first step in screening the data, we construct a null hypothesis based on the extreme assumption that, in each game considered separately, the probability of choosing  $s_3$  is the same for all players and is equal to the relative frequency observed in the experiment. Given this assumption, any observed heterogeneity of behaviour within any given game would be entirely due to random noise. We then find the actual number of  $s_3$  choices made by each subject in the eighteen games of the experiment and examine the cumulative frequency distribution of these numbers. We compare this observed distribution with the *benchmark* distribution implied by the null hypothesis, given the actual number of subjects in our sample.

Figure 2 plots the observed distribution and shows the 95 per cent confidence limits of the benchmark distribution. It is clear from inspection that the two distributions are different: the observed distribution has more weight in the tails, indicating that (as one would expect) there is some individual heterogeneity. Statistically, this difference is highly significant ( $\chi^2$

---

<sup>18</sup> Apart from one minor exception, level- $k$  theory implies that  $s_3$  is never chosen by players at level  $L1$  or above in any of these games. The exception is that in G5, P1 players at  $L1$  are predicted to choose  $s_3$  with probability  $1/3$ .



=195.35,  $p < 0.001$ ).<sup>19</sup> Nevertheless, the observed distribution is quite close to the benchmark distribution, with the same general S-shape, indicative of the absence of discrete subject types.<sup>20</sup>

*[Figure 2 near here]*

A similar analysis can be carried out separately for each of Conjectures 1, 2 and 3. Take the case of Conjecture 1. There are five pairs of games, namely (G1, G2), (G3, G4), (G5, G6), (G8, G9) and (G8, G10), for which Conjecture 1 implies that we should find more  $s_3$  choices in the second game than in the first. In any one of these pairs of games, a subject's choices are *in accordance with* Conjecture 1 if she does not choose  $s_3$  in the first game but does so in the second. We find the actual number of cases in which each subject behaves in accordance with Conjecture 1 and examine the cumulative frequency distribution of these numbers. We compare this observed distribution with the benchmark distribution implied by the null hypothesis, given the actual number of subjects in our sample. The results are shown in Figure 3a. Corresponding results for Conjectures 2 and 3 are shown in Figures 3b and 3c.

*[Figures 3a, 3b and 3c near here]*

In the case of Conjecture 1, we cannot reject the null hypothesis that observed heterogeneity is the result of random noise ( $\chi^2=4.08$ ,  $p = 0.395$ ). In the other two cases, there are significant differences between the observed and benchmark distributions (for Conjecture 2,  $\chi^2=14.91$ ,  $p = 0.011$ ; for Conjecture 3,  $\chi^2=14.98$ ,  $p = 0.004$ ). The observed distributions have more weight in their tails than the benchmark distributions do, indicating some degree of individual heterogeneity. Again, however, there is little indication of discrete subject types.<sup>21</sup>

---

<sup>19</sup> The confidence limits shown in Figure 2 were derived by simulation. For this, we generated 200 distributions with  $n=126$  by assigning to each hypothetical subject in each distribution a probability of choosing  $s_3$  in each game equal to the observed frequency of  $s_3$  choice in that particular game. Our chi-squared tests compare the observed distribution of relative frequencies of 'number of  $s_3$  choices' with the benchmark distribution.

<sup>20</sup> The slight irregularity at high values of 'number of  $s_3$  choices' suggests the possibility that a small minority of subjects (less than 10 per cent) might belong to a discrete type that always uses team reasoning.

<sup>21</sup> A further method of screening for subject heterogeneity is to look for patterns in the correlation between proportions of  $s_3$  choices in pairs of games. The matrix of correlation coefficients is reported in Appendix 5 [intended for online publication]. This matrix shows a general tendency for positive

### 3.6. *Learning*

It is also relevant to ask whether, over the course of the experiment, there was any significant trend in subjects' propensities to act according to the theory of team reasoning. Because the order in which games were played was randomized, it is sufficient to look for trends in the proportion of  $s_3$  choices. In fact, there is no significant trend.

## 4. Discussion

Our key findings concern cross-game variations in players' propensities to choose  $s_3$  that are inconsistent both with the theory of team reasoning and with level- $k$  theory. Specifically, we find that the frequency of  $s_3$  choices is influenced by the extent of payoff inequality in the  $s_1$ ,  $s_2$  and  $s_3$  equilibria. Equality of  $s_3$  payoffs and (possibly to a lesser extent) inequality of  $s_1$  and  $s_2$  payoffs tend to induce more  $s_3$  choices. In addition,  $s_3$  choices tend to be less frequent if, ex post, the payoffs in the  $s_3$  equilibrium are Pareto-dominated by those of other two equilibria than if the equilibria are not Pareto-comparable. We have argued that these effects are consistent with plausible conjectures about the conditions under which players are more likely to use one or other of the two modes of reasoning we have considered. Our provisional conclusion is that a satisfactory theory of behaviour in coordination games needs to include both collective optimality reasoning and bounded best response reasoning.

So far, however, we have shown only that the effects we have observed are inconsistent with the theory of team reasoning and with level- $k$  theory. Are we entitled to claim that this is true of the collective optimality approach *in general* and of bounded best response approach *in general*?

We need to consider a wide range of variant theories. To keep the exposition uncluttered, we do not present the predictions of all these theories for all the games in the experiment. Instead, we focus on four games, G1, G3, G7 and G11, which together encapsulate the problem of finding a single explanatory theory. For all these games, team reasoning theory predicts the choice of  $s_3$ , while level- $k$  theory predicts that  $s_3$  will not be chosen at any level above  $L0$ . The team reasoning prediction fails badly in G1 and G3, with

---

correlation, indicative of heterogeneity, but we do not discern any specific pattern to suggest that subjects can be divided into discrete types.

only 11.9 per cent and 13.5 per cent of players choosing  $s_3$ ; the level- $k$  prediction is similarly unsuccessful in G7 and G11, where 69.0 per cent and 88.9 per cent of players chose  $s_3$ . Similar failures of the two theories have been observed in previous experiments (see Section 1.3 above).

First, we ask whether any variant of the bounded best response approach could explain the choice of  $s_3$  in G7 and G11. Two obvious candidates are cognitive hierarchy theory and CGR's 'payoff and label salience' specification of  $L0$  behaviour (described in Section 1.2 above). But consider the condition (*Condition X*) that  $x > v, w$  (that is, for each player, the better payoff from  $s_1$  or  $s_2$  is strictly greater than the payoff from  $s_3$ ). Notice that Condition X holds for G7 and G11 (as it does for G1, G2, G3, G4, G6 and G9). It is easy to see that, given Condition X, cognitive hierarchy theory replicates the level- $k$  implication that  $s_3$  is not chosen by either player at any level above  $L0$ . Under the same condition,  $s_3$  is payoff-dominated for both players. Thus, even if  $s_3$  were label-salient, CGR's tie-breaking rule would never make it the modal choice at  $L0$  for either player, and so  $s_3$  would not be chosen at any level above  $L0$ . Notice also that if Condition X holds, the implications of level- $k$  theory are independent of attitudes to risk.

Another possibility is that players use level- $k$  reasoning but are averse to inequality, as in the theories of Fehr and Schmidt (1999) and Bolton and Ockenfels (2000), with the result that utility payoffs differ from material ones. But if inequality aversion is to induce  $L1$  reasoners (and thereby higher-level reasoners) to choose  $s_3$  in G7, we require that at least one player prefers the outcome in which both players receive a material payoff of 9 to the outcome in which she receives 10 and her co-player receives 9. This would be an extraordinarily high degree of aversion to advantageous inequality, implying a preference for unilaterally burning one's own money whenever one is better off than others – a possibility that Fehr and Schmidt rule out as 'very implausible' (p. 824).

More intuitively, it is hard to think of any plausible account of why players choose  $s_3$  in G7 and G11 which does not refer to players' perceptions of properties of strategy pairs as potential objects of coordination. The most obvious explanations of this behaviour use one or both of two facts – that the  $s_3$  equilibrium is the only pure-strategy equilibrium with equal payoffs, and that, because of isomorphism, the  $s_1$  and  $s_2$  equilibria are not attainable. But these are considerations that are relevant to players who are consciously trying to coordinate on some equilibrium, and not to players who are thinking only about best responses to one another's given strategy choices.

We now ask the opposite question: Is there any variant of the collective optimality approach that can explain why  $s_3$  is *not* chosen in G1 and G3? Recall that our specification of team reasoning theory already allows for inequality aversion and risk aversion. Within the collective optimality approach, the most obvious alternative to team reasoning is a theory which uses payoff dominance as a principle of equilibrium selection. Because of the isomorphism between  $s_1$  and  $s_2$ , the crucial issue in a payoff-dominance analysis is whether the pure-strategy  $s_3$  equilibrium gives higher expected utility to both players than the Nash equilibrium mix of  $s_1$  and  $s_2$ . In general, the Nash equilibrium mix of two isomorphic strategies can be different from the team-optimal mix, but this is not the case if, as in G1 and G3, both players are indifferent between the corresponding pure-strategy equilibria. In order for  $s_3$  to be payoff-dominant in G1 and G3, it is sufficient that the certainty of eight units of material payoff is preferred to a lottery which gives ten units with probability 0.5 and zero otherwise. This requires only that players are not extreme risk-seekers.

More intuitively, it seems obvious that if the  $s_3$  equilibrium is strictly Pareto-dominated by both the other pure-strategy equilibria, any plausible theory of bounded best-response reasoning will imply that players of G1 and G3 will not choose  $s_3$ . But this intuition cannot be carried over to collective-optimality reasoning unless players believe that one or other of the  $s_1$  and  $s_2$  equilibria is attainable. As we showed in Section 3.1, our subjects' use of the relative positions of the three circles is not consistent with their having that belief.

A final possibility (mentioned in Section 1.2 above) is that the population of players is a mixture of naïve and sophisticated team reasoners. Consider a model in which naïve players do not recognise the difficulty of coordinating on one of two isomorphic equilibria, and simply choose any strategy associated with an equilibrium that is team-optimal *ex post*. Such players would choose  $s_1$  or  $s_2$  in G1 and G3. Sophisticated players who attached a sufficiently high probability to their co-players being naïve would then avoid  $s_3$ , despite its team-optimality. A model of this kind could organise the main patterns in our data. However, as explained in Section 1.3, it would not explain the tendency, observed in other experiments, for conflicts of interest to reduce the power of focal points.

## 5. Conclusion

Game theorists have proposed two very different ways of modelling players' reasoning in coordination games. One approach, exemplified by level- $k$  theory, assumes boundedly rational best-response reasoning by individual players. An alternative approach, exemplified

by the theory of team reasoning, assumes that players select strategies by using some criterion of collective optimality. Previous authors have speculated that both approaches may play some role in explaining observed behaviour, but we believe that our paper is the first to translate these speculations into testable conjectures about which properties of coordination games facilitate or inhibit these two modes of reasoning and then to carry out the tests.

Our experimental results support the hypothesis that both modes of reasoning are used in coordination games, and that which mode is more likely to be used depends on the particular characteristics of the game being played. Considered alongside the findings of previous experiments, our results suggest that a one-theory-fits-all approach is unlikely to provide a satisfactory explanation of behaviour in coordination games. What is needed is a theory of how the mode of reasoning that players use is influenced by properties of the particular coordination game they are playing. The conjectures we have proposed in this paper, and the supporting experimental evidence we have reported, are first steps in this enterprise.

**Table 1: Games used in the experiment**

Payoff (in)equality	$s_3$ strictly Pareto-dominated $y > v$	$s_3$ weakly Pareto-dominated $y = v$ and $(x > y$ or $v > w)$	$s_3$ not Pareto-comparable $v > y$ and $x > w$	$s_3$ weakly Pareto-dominating $w = x$ and $(x > y$ or $v > w)$	$s_3$ strictly Pareto-dominating $w > x$
$x = y, v > w$	G1= $\langle(10,10), (10,10), (9,8)\rangle$	G5= $\langle(10,10), (10,10), (10,9)\rangle$	G8= $\langle(10,10), (10,10), (11,9)\rangle$	G12= $\langle(10,10), (10,10), (11,10)\rangle$	G15= $\langle(10,10), (10,10), (12,11)\rangle$
$x > y, v > w$	G2= $\langle(10,9), (9,10), (8,7)\rangle$	G6= $\langle(10,9), (9,10), (9,8)\rangle$	G9= $\langle(12,9), (9,12), (11,10)\rangle$ G10= $\langle(11,10), (10,11), (12,9)\rangle$	G13= $\langle(10,9), (9,10), (11,10)\rangle$	G16= $\langle(10,9), (9,10), (12,11)\rangle$
$x = y, v = w$	G3= $\langle(10,10), (10,10), (9,9)\rangle$	n.a.	n.a.	n.a.	G17= $\langle(10,10), (10,10), (11,11)\rangle$
$x > y, v = w$	G4= $\langle(10,9), (9,10), (8,8)\rangle$	G7= $\langle(10,9), (9,10), (9,9)\rangle$	G11= $\langle(11,9), (9,11), (10,10)\rangle$	G14= $\langle(10,9), (9,10), (10,10)\rangle$	G18= $\langle(10,9), (9,10), (11,11)\rangle$

**Table 2: Predicted proportions of  $s_3$  choices**

Proportion of  $s_3$  choices predicted by:

---

Game	Team reasoning P1 and P2	Level- $k$ theory for level and player:							
		$L0$		$L1$		$L2$		$L3, L4, \dots$	
		P1	P2	P1	P2	P1	P2	P1	P2
G1	1	1/3	1/3	0	0	0	0	0	0
G2	1	1/3	1/3	0	0	0	0	0	0
G3	1	1/3	1/3	0	0	0	0	0	0
G4	1	1/3	1/3	0	0	0	0	0	0
G5	1	1/3	1/3	1/3	0	0	0	0	0
G6	1	1/3	1/3	0	0	0	0	0	0
G7	1	1/3	1/3	0	0	0	0	0	0
G8	1	1/3	1/3	1	0	0	1	1	0
G9	1	1/3	1/3	0	0	0	0	0	0
G10	1	1/3	1/3	1	0	0	1	1	0
G11	1	1/3	1/3	0	0	0	0	0	0
G12	1	1/3	1/3	1	1/3	1	1	1	1
G13	1	1/3	1/3	1	1/2	1	1	1	1
G14	1	1/3	1/3	1/2	1/2	1	1	1	1
G15	1	1/3	1/3	1	1	1	1	1	1
G16	1	1/3	1/3	1	1	1	1	1	1
G17	1	1/3	1/3	1	1	1	1	1	1
G18	1	1/3	1/3	1	1	1	1	1	1

---

**Table 3: Distribution of choices between  $s_1$  and  $s_2$** 

distribution of choices over ( $s_1, s_2$ )

---

game	treatment A (left, right)	treatment B (bottom, left)	treatment C (right, bottom)	all treatments (not reported if $s_1, s_2$ isomorphic)
G1	(25, 12)**	(13, 25)*	(17, 19)	
G2†	(21, 20)	(22, 19)	(19, 23)	(62, 62)
G3	(21, 14)	(12, 24)*	(24, 14)	
G4	(12, 10)	(10, 12)	(6, 21)***	
G5	(27, 7)***	(16, 21)	(18, 18)	
G6†	(24, 15)	(17, 25)	(19, 19)	(60, 59)
G7	(8, 8)	(5, 6)	(4, 8)	
G8	(26, 8)***	(10, 25)**	(18, 17)	
G9†	(8, 6)	(5, 10)	(2, 5)	(15, 21)
G10†	(21, 10)*	(17, 19)	(14, 20)	(52, 49)
G11	(2, 1)	(3, 3)	(2, 3)	
G12	(11, 5)	(11, 15)	(16, 6)*	
G13†	(2, 2)	(3, 4)	(0, 5)*	(5, 11)
G14	(0, 0)	(0, 2)	(0, 1)	
G15	(7, 6)	(5, 5)	(9, 4)	
G16†	(2, 2)	(0, 3)	(4, 4)	(6, 9)
G17	(1, 0)	(1, 2)	(0, 1)	
G18	(0, 1)	(0, 2)	(0, 1)	

---

*Note:* Games in which  $s_1$  and  $s_2$  are non-isomorphic, allowing meaningful ‘all treatments’ tests, are marked by †. The total number of subjects (including those who chose  $s_3$ ) in each treatment was 42. Asterisks report two-sided binomial tests of the hypothesis that  $s_1$  and  $s_2$  are chosen with equal probability; \*, \*\* and \*\*\* denote rejection at 10, 5 and 1 per cent significance levels.



**Table 4: Observed percentages of  $s_3$  choices**

payoff (in)equality	$s_3$ strictly Pareto-dominated $y > v$	$s_3$ weakly Pareto-dominated $y = v$ and ( $x > y$ or $v > w$ )	$s_3$ not Pareto- comparable $v > y$ and $x > w$	$s_3$ weakly Pareto- Pareto-dominating $w = x$ and ( $x > y$ or $y > w$ )	$s_3$ strictly Pareto-dominating $w > x$
$x = y, v > w$	11.9### (G1)	15.1### (G5)	17.5### (G8)	49.2*** (G12)	71.4*** (G15)
$x > y, v > w$	1.6### (G2)	5.6### (G6)	69.0*** (G9) 19.8## (G10)	87.3*** (G13)	88.1*** (G16)
$x = y, v = w$	13.5### (G3)				96.0*** (G17)
$x > y, v = w$	43.7** (G4)	69.0*** (G7)	88.9*** (G11)	97.6*** (G14)	96.8*** (G18)

*Note:* For all games,  $n = 126$ . Asterisks and hashes report two-sided binomial tests of the null hypothesis that  $s_3$  is chosen with probability  $1/3$ .

\*, \*\* and \*\*\* denote rejection at 10, 5 and 1 per cent significance levels when observed proportion is greater than  $1/3$ . #, ## and ### denote rejection at 10, 5 and 1 per cent significance levels when observed proportion is less than  $1/3$ .

**Table 5: Tests of Hypotheses 3.1–3.3**

game pair	test of	effect predicted by level- $k$ theory	percentages of $s_3$ choices	$\chi^2$
G1, G2	Hypothesis 3.1	no change	(11.9, 1.6)	11.27###
G3, G4	Hypothesis 3.1	no change	(13.5, 43.7)	28.88***
G5, G6	Hypothesis 3.1	decrease	(15.1, 5.6)	6.55##
G8, G9	Hypothesis 3.1	decrease	(17.5, 69.0)	49.71***
G8, G10	Hypothesis 3.1	no change	(17.5, 19.8)	0.29
G1, G3	Hypothesis 3.2	no change	(11.9, 13.5)	0.22
G2, G4	Hypothesis 3.2	no change	(1.6, 43.7)	51.07***
G6, G7	Hypothesis 3.2	no change	(5.6, 69.0)	78.05***
G9, G11	Hypothesis 3.2	no change	(69.0, 88.9)	21.55***
G10, G11	Hypothesis 3.2	decrease	(19.8, 88.9)	76.45***
G13, G14	Hypothesis 3.2	decrease	(87.3, 97.6)	8.89***
G2, G9	Hypothesis 3.3	no change	(1.6, 69.0)	85.00***
G6, G9	Hypothesis 3.3	no change	(5.6, 69.0)	74.42***
G4, G11	Hypothesis 3.3	no change	(43.7, 88.9)	51.57***
G7, G11	Hypothesis 3.3	no change	(69.0, 88.9)	18.94**

*Note:*  $\chi^2$ -statistics for McNemar test for paired  $s_3$  proportions. \*, \*\* and \*\*\* denote rejection at 10, 5 and 1 per cent significance levels when observed effect is an increase. #, ## and ### denote rejection at 10, 5 and 1 per cent significance levels when observed effect is a decrease. Paired  $t$ -tests on differences between  $s_3$  proportions give the same results.

**Table 6: Regression analysis**

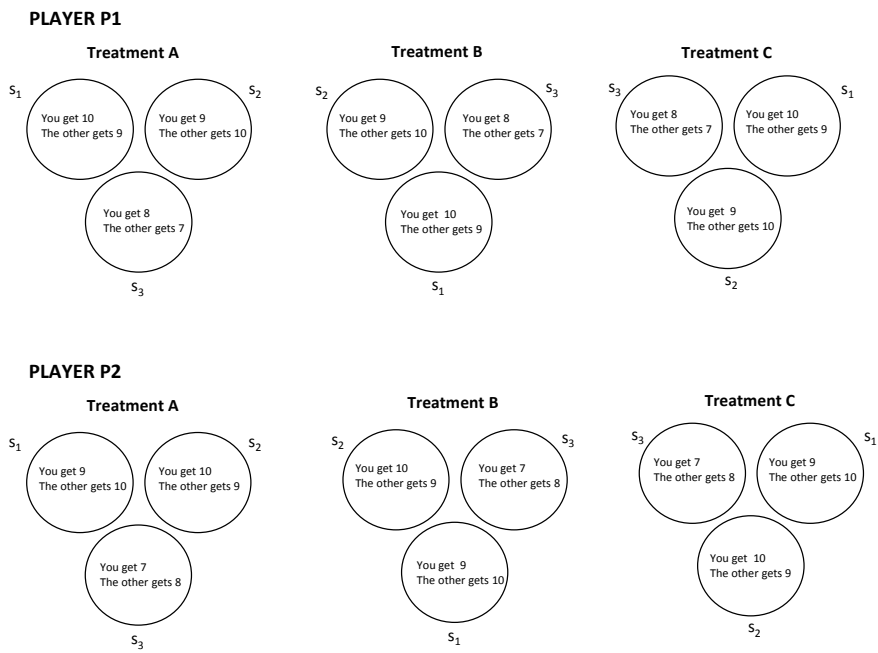
	Random effect probit coefficients (standard errors)	Average marginal effects
<i>x equal to y</i>	-0.56*** (0.78)	-0.13
<i>v equal to w</i>	1.13*** (0.08)	0.26
Pareto dominating ( $\beta_{PDing}$ )	0.65*** (0.12)	0.15
Pareto dominated ( $\beta_{PDed}$ )	-0.35*** (0.13)	-0.81
Player P1	0.01 (0.07)	0.03
<i>s</i> <sub>3</sub> left	-0.11 (0.08)	-0.02
<i>s</i> <sub>3</sub> right	-0.17* (0.09)	-0.04
Average payoff ratio	-5.43*** (0.80)	-1.24
Experiments	-0.01* (0.01)	-0.003
Economics	0.07 (0.10)	0.01
Nationality	0.35* (0.18)	0.08
Gender	-0.22** (0.10)	-0.05
Age	-0.03 (0.02)	-0.007
Constant	5.87 (0.98)	
$\beta_{PDing} - \beta_{PDed}$	1.00*** (0.20)	

$n= 1944$ ; Log likelihood = -823.02; Wald  $\chi^2=603.46$

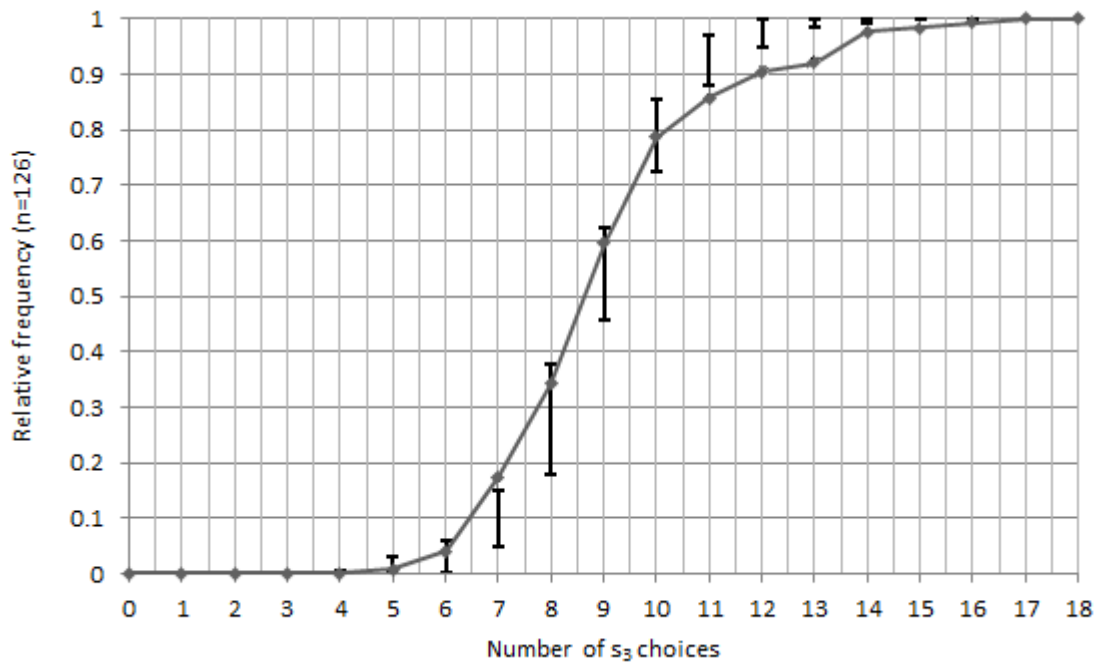
The dependent variable takes value 1 if the choice is the one predicted by team reasoning (*s*<sub>3</sub>) theory and 0 otherwise. *Player P1*: dummy variable taking value 1 if the subject plays as P1; *x equal to y*: dummy variable taking value 1 if the subject plays a game in which  $x=y$ ; *v equal to w*: dummy variable taking value 1 if the subject plays a game in which  $v=w$ ; *Pareto dominated*: dummy variable taking value 1 if the subject plays a game in which *s*<sub>3</sub> is Pareto dominated (weakly or strongly) ex-post, by *s*<sub>1</sub> and *s*<sub>2</sub>; *Pareto dominating*: dummy variable taking value 1 if the subject plays a game in which *s*<sub>3</sub> Pareto-dominates (weakly or strongly), ex-post, *s*<sub>1</sub> and *s*<sub>2</sub>; *s*<sub>3</sub> left: dummy variable taking value 1 if the subject plays a game in which *s*<sub>3</sub> is in the left circle; *s*<sub>3</sub> right: dummy variable taking value 1 if the subject plays a game in which *s*<sub>3</sub> is in the right circle; *Average payoff ratio*: ratio between the average of P1's and P2's payoffs in *s*<sub>1</sub> and the average of their payoffs in *s*<sub>3</sub>; *Experiments*: number of previous experiments the subjects has participated in; *Economics*: dummy variable taking value 1 if the subject is enrolled in economics/management university courses; *Nationality*: dummy variable taking value 1 if the subject is Italian; *Gender*: dummy variable taking value 1 if the subject is female; *Age*: subject's age.

\*\*\* significant at 1%; \*\* significant at 5%; \* significant at 10%

**Figure 1: Payoff displays for a typical game (G2)**



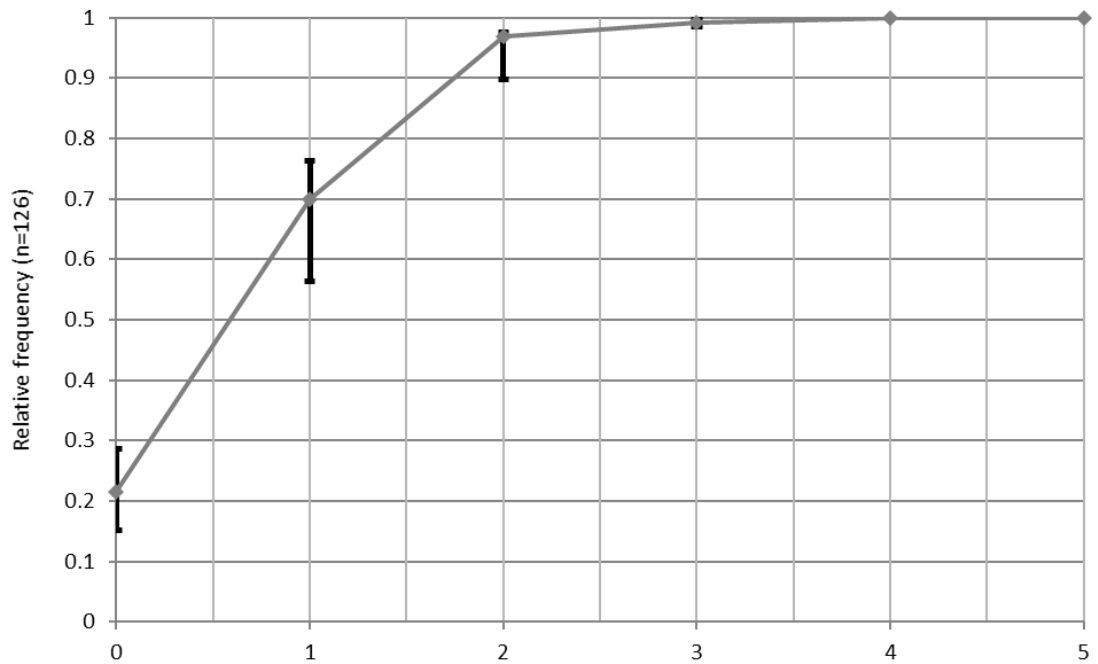
**Figure 2: Relative frequency of number of  $s_3$  choices in games G1-G18**



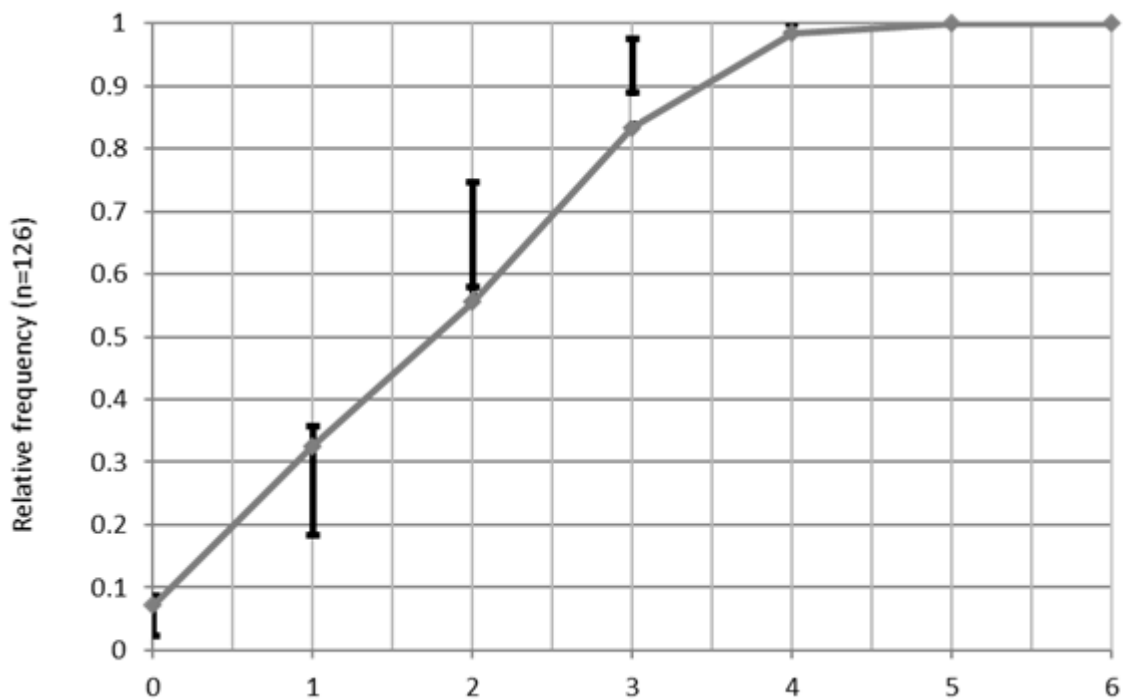
The graph shows the cumulative relative frequency of the number of  $s_3$  choices in the eighteen games of the experiment. The bars show the 95 per cent confidence limits of the benchmark distribution.

**Figure 3: Relative frequency of number of choices in accordance with Conjectures 1, 2 and 3**

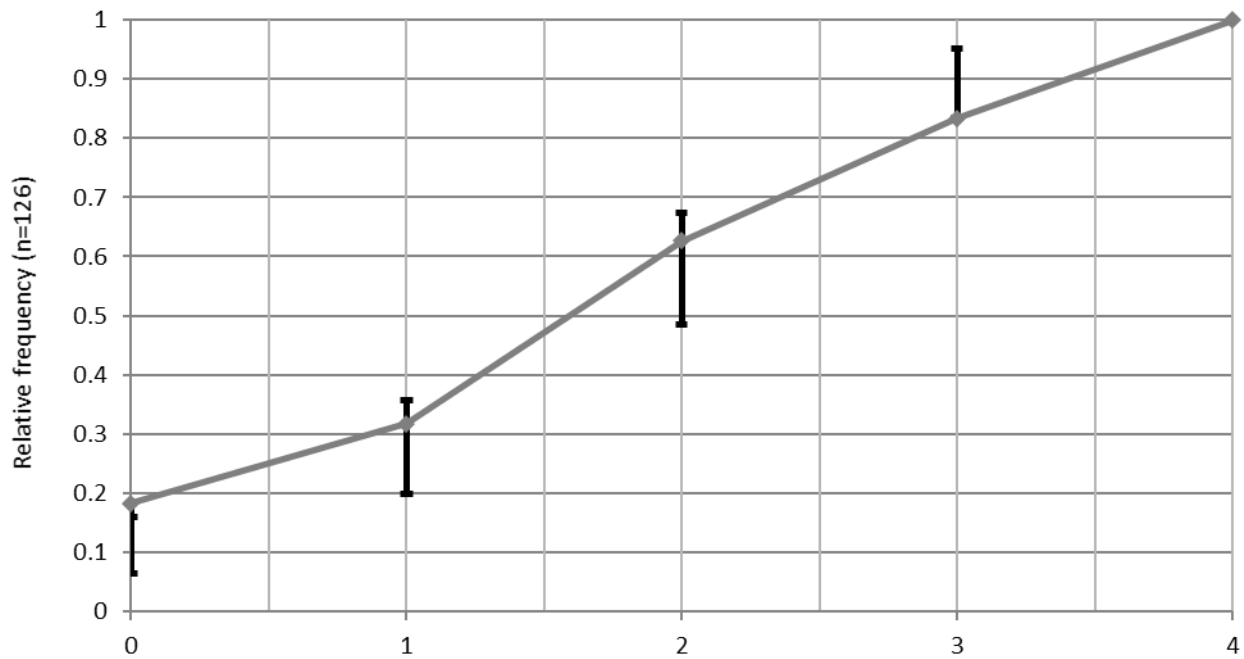
*a: Choices in accordance with Conjecture 1*



*b: Choices in accordance with Conjecture 2*



*c: Choices in accordance with Conjecture 3*



In each panel the graph shows the cumulative relative frequency of the number of choices in accordance with the relevant conjecture. The bars show the 95 per cent confidence limits of the benchmark distribution.

## References

- Bacharach, Michael (1993). Variable universe games. In Ken Binmore, Alan Kirman, and Piero Tani (eds) *Frontiers of Game Theory*. Cambridge, MA: MIT Press. pp. 255–275.
- Bacharach, Michael (1999). Interactive team reasoning: a contribution to the theory of cooperation. *Research in Economics* 53: 117–147.
- Bacharach, Michael (2006). *Beyond Individual Choice: Teams and Frames in Game Theory*. Princeton, NJ: Princeton University Press. Edited by Natalie Gold and Robert Sugden.
- Bacharach, Michael and Michele Bernasconi (1997). The variable frame theory of focal points: an experimental study. *Games and Economic Behavior* 19: 1–45.
- Bacharach, Michael and Dale O. Stahl (2000). Variable-frame level- $n$  theory. *Games and Economic Behavior* 33(2): 220–246.
- Bardsley, Nicholas, Judith Mehta, Chris Starmer and Robert Sugden (2010). Explaining focal points: cognitive hierarchy theory *versus* team reasoning. *Economic Journal* 120: 40–79.
- Bardsley, Nicholas and Aljaž Ule (2017). Focal points revisited: team reasoning, the principle of insufficient reason and cognitive hierarchy theory. *Journal of Economic Behavior and Organization* 133: 74–86.
- Bett, Zoë, Anders Poulsen and Odile Poulsen (2016). The focality of dominated compromises in tacit coordination situations: experimental evidence. *Journal of Behavioral and Experimental Economics* 60: 29–34.
- Blume, Andreas (2000). Coordination and learning with a partial language. *Journal of Economic Theory* 95: 1–36.
- Blume, Andreas and Uri Gneezy (2010). Cognitive forward induction and coordination without common knowledge: an experimental study. *Games and Economic Behavior* 68: 488–511.
- Bolton, Gary and Axel Ockenfels (2000). ERC: a theory of equity, reciprocity and competition. *American Economic Review* 90: 166–193.



- Camerer, Colin F., Teck-Hua Ho and Juin-Kuan Chong (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics* 119: 861–98.
- Casajus, André (2000). Focal points in framed strategic forms. *Games and Economic Behavior* 32: 263-291.
- Colman, Andrew M., Briony D. Pulford and Jo Rose (2008). Collective rationality in interactive decisions: evidence for team reasoning. *Acta Psychologica* 128: 387–397.
- Colman, Andrew M., Briony D. Pulford and Catherine L. Lawrence (2014). Explaining strategic coordination: cognitive hierarchy theory, strong Stackelberg reasoning, and team reasoning. *Decision* 1: 35–38.
- Costa-Gomes, Miguel A., Vincent P. Crawford, and Bruno Broseta (2001). Cognition and behavior in normal-form games: an experimental study. *Econometrica* 69(5): 1193–1235.
- Crawford, Vincent P., Miguel A. Costa-Gomes and Nagore Iriberry (2013). Structural models of nonequilibrium strategic thinking: theory, evidence, and applications. *Journal of Economic Literature* 51: 5–62.
- Crawford, Vincent P., Uri Gneezy and Yuval Rottenstreich (2008). The power of focal points is limited: even minute payoff asymmetry may yield large coordination failures. *American Economic Review*, 98: 1443–1458.
- Crawford, Vincent P. and Hans Haller (1990). Learning how to cooperate: optimal play in repeated coordination games. *Econometrica* 58: 571–595.
- Fehr, Ernst and Klaus Schmidt (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics* 114: 817–868.
- Fischbacher, Urs (2007). z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental Economics* 10): 171–78.
- Hargreaves Heap, Shaun, David Rojo Arjona and Robert Sugden (2014). How portable is level-0 behaviour? A test of level- $k$  theory in games with non-neutral frames. *Econometrica* 82: 1133–1151.
- Harsanyi, John and Reinhard Selten (1988). *A General Theory of Equilibrium Selection in Games*. Cambridge, MA: The MIT Press.
- Hodgson, David H. (1967). *Consequences of Utilitarianism*. Oxford: Clarendon Press.

- Isoni, Andrea, Anders Poulsen, Robert Sugden and Kei Tsutsui (2013). Focal points in tacit bargaining games: experimental evidence. *European Economic Review* 59 (2013): 167–188.
- Lewis, David K. (1969). *Convention: A Philosophical Study*, Cambridge, MA: Harvard University Press.
- Mehta, Judith, Chris Starmer and Robert Sugden (1994). The nature of salience: an experimental investigation of pure coordination games. *American Economic Review*, 84: 658–73.
- Nagel, Rosemarie (1995). Unraveling in guessing games: an experimental study. *American Economic Review* 85: 1313–26.
- Regan, Donald (1980). *Utilitarianism and Cooperation*. Oxford: Clarendon Press.
- Rubinstein, Ariel, Amos Tversky and Dana Heller (1996). Naïve strategies in competitive games. In *Understanding Strategic Interaction: Essays in Honor of Reinhard Selten*, ed. Wulf Albers, Werner Güth, Peter Hammerstein, Benny Moldovanu and Eric van Damme, 394–402. Berlin: Springer-Verlag.
- Schelling, Thomas C. (1960). *The Strategy of Conflict*. Cambridge, MA: Harvard University Press.
- Smerilli, Alessandra (2012). We-thinking and vacillation between frames: filling a gap in Bacharach's theory. *Theory and Decision* 73: 539–560.
- Stahl, Dale O. and Paul W. Wilson (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior and Organization* 25: 309–327.
- Sugden, Robert (1993). Thinking as a team: toward an explanation of nonselfish behavior. *Social Philosophy and Policy* 10: 69–89.
- Sugden, Robert (1995). A theory of focal points. *Economic Journal* 105: 533–550.
- Sugden, Robert (2008). Nash equilibrium, team reasoning and cognitive hierarchy theory. *Acta Psychologica* 128: 402–404.
- Tan, Jonathan and Daniel Zizzo (2008). Groups, cooperation and conflict in games. *Journal of Socio-Economics* 37: 1–17.

van Elten, Jonas and Stephan Penczynski (2015). Coordination games with asymmetric payoffs: an experimental study with intra-group communication. Working paper available at <http://www.penczynski.de/attach/APC.pdf>.

**FOR ONLINE PUBLICATION ONLY**

**Appendix 1: Predictions of level- $k$  theory**

Table A1.1 shows the predictions of level- $k$  theory for each of nine cases, defined by the relative values of the parameters  $v$ ,  $w$ ,  $x$  and  $y$ , under the assumption that  $L0$  choices are uniformly random. Although predictions are shown only for levels up to  $L3$ , the predictions for  $L4$ ,  $L6$ , ... are always the same as for  $L2$  and those for  $L5$ ,  $L7$ , ... are always the same as for  $L3$ . These cases are not exhaustive, but they include all our games G1–G18. Given the assumption that  $L0$  choices are random, and given the assumption that whenever two or more strategies have the same expected payoff they are chosen with equal probability, the derivation of the figures in the table is straightforward.

**Table A1.1: Level- $k$  predictions**

game(s)	player	proportion of choices that are $(s_1, s_2, s_3)$			
		$L0$	$L1$	$L2$	$L3$
Case 1: $x = y > v$					
(G1, G3)					
	P1	(1/3, 1/3, 1/3)	(1/2, 1/2, 0)	(1/2, 1/2, 0)	(1/2, 1/2, 0)
	P2	(1/3, 1/3, 1/3)	(1/2, 1/2, 0)	(1/2, 1/2, 0)	(1/2, 1/2, 0)
Case 2: $x > v, x > y$					
(G2, G4, G6, G7, G9, G11)					
	P1	(1/3, 1/3, 1/3)	(1, 0, 0)	(0, 1, 0)	(1, 0, 0)
	P2	(1/3, 1/3, 1/3)	(0, 1, 0)	(1, 0, 0)	(0, 1, 0)
Case 3: $x = y = v > w$					
(G5)					
	P1	(1/3, 1/3, 1/3)	(1/3, 1/3, 1/3)	(1/2, 1/2, 0)	(1/2, 1/2, 0)
	P2	(1/3, 1/3, 1/3)	(1/2, 1/2, 0)	(1/2, 1/2, 0)	(1/2, 1/2, 0)
Case 4: $v > x = y > w$					
(G8)					
	P1	(1/3, 1/3, 1/3)	(0, 0, 1)	(1/2, 1/2, 0)	(0, 0, 1)

	P2	$(1/3, 1/3, 1/3)$	$(1/2, 1/2, 0)$	$(0, 0, 1)$	$(1/2, 1/2, 0)$
Case 5: $v > x > y > w$					
(G10)					
	P1	$(1/3, 1/3, 1/3)$	$(0, 0, 1)$	$(0, 1, 0)$	$(0, 0, 1)$
	P2	$(1/3, 1/3, 1/3)$	$(0, 1, 0)$	$(0, 0, 1)$	$(0, 1, 0)$
Case 6: $v > x = w = y$					
(G12)					
	P1	$(1/3, 1/3, 1/3)$	$(0, 0, 1)$	$(0, 0, 1)$	$(0, 0, 1)$
	P2	$(1/3, 1/3, 1/3)$	$(1/3, 1/3, 1/3)$	$(0, 0, 1)$	$(0, 0, 1)$
Case 7: $v > x = w > y$					
(G13)					
	P1	$(1/3, 1/3, 1/3)$	$(0, 0, 1)$	$(0, 0, 1)$	$(0, 0, 1)$
	P2	$(1/3, 1/3, 1/3)$	$(0, 1/2, 1/2)$	$(0, 0, 1)$	$(0, 0, 1)$
Case 8: $x = v = w > y$					
(G14)					
	P1	$(1/3, 1/3, 1/3)$	$(1/2, 0, 1/2)$	$(0, 0, 1)$	$(0, 0, 1)$
	P2	$(1/3, 1/3, 1/3)$	$(0, 1/2, 1/2)$	$(0, 0, 1)$	$(0, 0, 1)$
Case 9: $v \geq w > x$					
(G15, G16, G17, G18)					
	P1	$(1/3, 1/3, 1/3)$	$(0, 0, 1)$	$(0, 0, 1)$	$(0, 0, 1)$
	P2	$(1/3, 1/3, 1/3)$	$(0, 0, 1)$	$(0, 0, 1)$	$(0, 0, 1)$

---

## FOR ONLINE PUBLICATION ONLY

### Appendix 2: Instructions

Buongiorno.

Ti ringraziamo per aver voluto prendere parte a questo esperimento sui processi decisionali.

L'esperimento durerà all'incirca 40 minuti.

Riceverai 3 euro per la partecipazione e nel corso dell'esperimento potrai guadagnare un'ulteriore somma che dipenderà dalle tue scelte e dalle scelte degli altri partecipanti.


Durante l'esperimento ti chiediamo di non comunicare con gli altri partecipanti.

Le risposte che fornirai e le scelte che farai saranno assolutamente anonime. Gli sperimentatori non saranno in grado di associare le tue scelte e le tue risposte al tuo nome.

Ti chiediamo di seguire attentamente le istruzioni che compariranno sul tuo schermo e che lo sperimentatore leggerà ad alta voce.

Se hai dubbi o vuoi ricevere chiarimenti su alcuni aspetti dell'esperimento alza la mano, uno degli sperimentatori sarà a tua disposizione.

Premi continua per proseguire.



Welcome to this experiment on decisional processes and thank you for participating in it.

The experiment will last approximately 40 minutes.

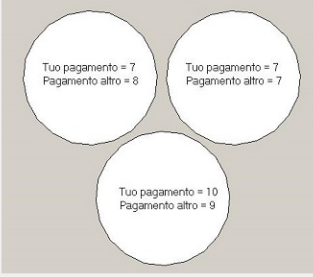
You will receive 3 euro for your participation. You can gain more money, depending on your choices and on choices of other participants.

Your answers and your choices will be completely anonymous. The experimenters are not able to associate your choices and your answers to your name.

We ask you to pay attention to the instructions that will appear on your screen. They will be read aloud by one of the experimenters.

If you have any doubts or questions about anything related to the experiment raise your hand: one of the experimenters will come.

Click on 'continue' to proceed.



LA SCELTA TIPO

Nel corso dell'esperimento dovrai prendere delle decisioni le cui caratteristiche fondamentali possono essere descritte attraverso l'illustrazione di una "scelta tipo".

Innanzitutto sarai accoppiato a un'altra persona presente in aula, ma di cui non ti sarà rivelata l'identità, così come all'altra persona non sarà rivelata la tua identità.

Sul monitor del tuo computer apparirà una figura come quella sopra. I numeri riportati nella figura, che sono puramente esemplificativi, rappresentano i pagamenti in euro per ciascuna combinazione di scelte.

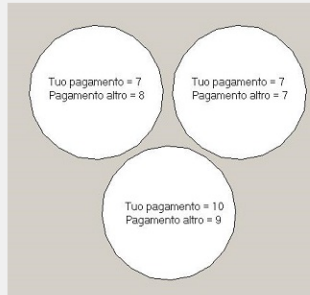
Premi continua per proseguire.

During the experiment you will be asked to make choices. The main characteristics of the choices can be described through an example.

First of all you will be matched with another person in this room. You will never know the identity of the other person, nor will he/she know your identity.

On the screen of your computer will appear a figure similar to the one you are now seeing above. The numbers on the figure, which are only examples, represent the payments in euro for each combination of choices.

Click on 'continue' to proceed.



Tu dovrai scegliere una delle tre opzioni cliccando su uno dei tre cerchi della figura.

La persona a cui sei stato accoppiato dovrà fare lo stesso.

Se scegli un cerchio che non corrisponde a quello scelto dall'altra persona entrambi otterrete un pagamento di 0 euro.

Se invece entrambi scegliete lo stesso cerchio, ognuno otterrà il pagamento che appare in quel cerchio.

Premi continua per proseguire.

Continua

You must choose one of the three options by clicking on one of the three circles on the figure.


The person you are matched with must do the same.

If your choice does not correspond to the choice of the other person, you will both get 0 euro.

If both of you choose the same option, each of you will obtain the payment which is in the circle.

Click on 'continue' to proceed.





Così, per esempio, se entrambi scegliete il cerchio in alto a sinistra nella figura, tu otterrai 7 euro e l'altro partecipante otterrà 8 euro.

Se entrambi scegliete il cerchio in alto a destra, sia tu che l'altro partecipante otterrete 7 euro.

Se invece entrambi scegliete il cerchio in basso, tu otterrai un pagamento di 10 euro, mentre l'altro partecipante otterrà un pagamento di 9 euro.

Se scegliere cerchi differenti, otterrete entrambi 0 euro.

Clicca continua per proseguire.

So, if both of you choose the top left circle, you will get 7 euro and the other will get 8 euro.

If both of you choose the top right circle, you will get 7 euro and the other will get 7 euro.

If both of you choose the bottom circle, you will obtain 10 euro and the other will get 7 euro.

If you choose different circles, both of you will get 0 euro.

Click on 'continue' to proceed.

#### L'ESPERIMENTO

L'esperimento consiste di 18 ripetizioni che chiameremo round. In ogni round dovrai prendere delle decisioni utilizzando diagrammi di scelta simili a quelli della decisione tipo.

All'inizio di ogni round sarai accoppiato a un altro partecipante e ti verrà presentato un diverso diagramma a cui sono associati pagamenti differenti.

Nota che in ogni round sarai accoppiato a una persona diversa.

Come nell'esempio, se tu e la persona a cui sei stato accoppiato scegliete lo stesso cerchio otterrete i pagamenti riportati in quel cerchio. Se invece i cerchi da voi scelti saranno diversi sia tu che l'altro partecipante otterrete zero euro.

Alla fine del round non riceverai alcuna informazione sul cerchio scelto dall'altra persona.

Al termine dell'esperimento, uno dei round verrà estratto a sorte, sarai allora informato sul cerchio scelto dall'altra persona in quel round. Se in quel round tu e l'altra persona avete scelto lo stesso cerchio, allora otterrete i pagamenti riportati in quel cerchio. Se tu e l'altra persona in quel round avete scelto due cerchi differenti allora otterrete zero euro.

In particolare, prima di iniziare l'esperimento, uno dei partecipanti estrarrà un busta da un contenitore contenente 18 buste. Ogni busta contiene un biglietto numerato da 1 a 18. Ogni busta contiene un solo biglietto.

La busta sarà consegnata allo sperimentatore, che l'aprirà alla fine dell'esperimento.

Questo numero corrisponderà al round scelto per il pagamento.

Clicca continua per proseguire.

Continua

The experiment.

The experiment consists of 18 rounds. In each round you must make decisions on figures similar to the ones you have seen before.

At the beginning of each round you will be matched with a person, and you will see a figure with different payments.

Notice that in each round you will be matched with a different person.

As in the previous example, if you and the other person choose the same circle, you will get the corresponding amount. If you and the other person choose a different circle the payment will be 0 euro.

At the end of each round you will not receive any feedback on the results.

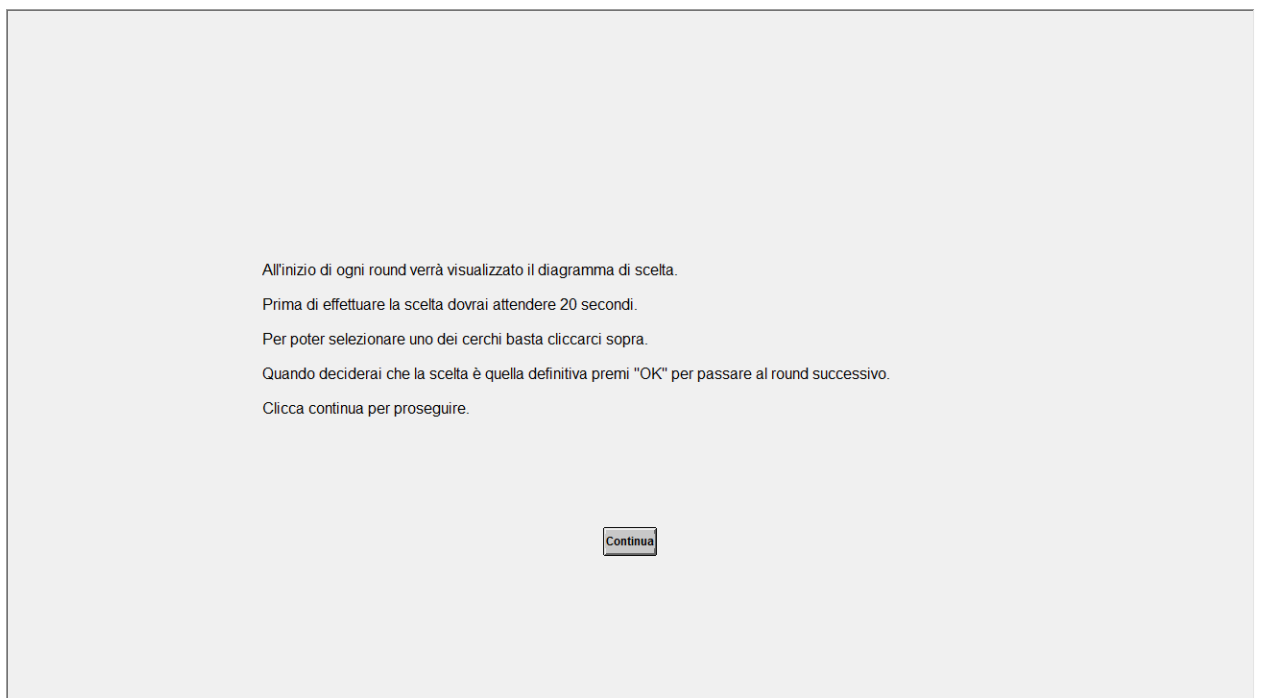
At the end of the experiment, one of the rounds will be randomly selected, and you will receive the corresponding payment associated to that round.

In particular before starting the experiment one participant will draw a ticket from a box containing 18 tickets numbered from 1 to 18.

The ticket (without being opened) will be given to the experimenter.

At the end of the experiment you will know the selected round.

Click on 'continue' to proceed.



At the beginning of each round the figure with three circles will be shown.

Before making your choice you must wait for 20 seconds.

In order to select a choice it is sufficient to click on the selected circle. If you want to change your choice, you must click again on the selected choice in order to deselect it, then you can click on the other choice.

When your decision is definitive you must click on the 'ok' button, in order to go to the following round.

Click on 'continue' to proceed.

Prima di iniziare con l'esperimento vero e proprio, parteciperai a una sessione di prova di tre round,  
In ogni round ti sarà chiesto di scegliere un'opzione e di rispondere a una domanda che trovi nel foglio  
"Domande di controllo".  
Le scelte che farai in questa sessione di prova non avranno alcun effetto sul pagamento finale.  
Alla fine della sessione di prova correggeremo le domande, e quando tutti i dubbi saranno chiariti  
procederemo con l'esperimento.  
Clicca continua per proseguire.

Continua

Before starting the experiment, you will participate in a trial session with 3 rounds.

In each round you will be asked to choose an option and to answer a question, which is on the sheet on your table.

The choices made in this session will not affect your final payment.

At the end of this session we shall correct the questions and we shall clarify any doubts. After that we shall proceed with the experiment.

Click on 'continue' to proceed.

## FOR ONLINE PUBLICATION ONLY

### Appendix 3: Matching procedure

The following tables report the details of the procedure used to match subjects and to define the sequences of games played by each subject in the 18 rounds. We start with the procedure used in SESSION 1. Columns 1-18 are subjects, rows 1-18 are rounds. The entry in each cell is the number of the game (G1–G18) that is played by the relevant subject in the relevant round. In each game, in each round, one subject from 1 to 9 is paired with one from 10 to 18.

#### STEP 1:

We assigned games 1 – 18 to player 10, using a random sequence :

round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1										11								
2										8								
3										1								
4										3								
5										6								
6										17								
7										4								
8										10								
9										7								
10										16								
11										14								
12										12								
13										15								
14										18								
15										5								
16										13								
17										9								
18										2								

## STEP 2:

We used the same random sequence to assign games to players 11-18 (with a shift for each column):

round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1										11	2	9	13	5	18	15	12	14
2										8	11	2	9	13	5	18	15	12
3										1	8	11	2	9	13	5	18	15
4										3	1	8	11	2	9	13	5	18
5										6	3	1	8	11	2	9	13	5
6										17	6	3	1	8	11	2	9	13
7										4	17	6	3	1	8	11	2	9
8										10	4	17	6	3	1	8	11	2
9										7	10	4	17	6	3	1	8	11
10										16	7	10	4	17	6	3	1	8
11										14	16	7	10	4	17	6	3	1
12										12	14	16	7	10	4	17	6	3
13										15	12	14	16	7	10	4	17	6
14										18	15	12	14	16	7	10	4	17
15										5	18	15	12	14	16	7	10	4
16										13	5	18	15	12	14	16	7	10
17										9	13	5	18	15	12	14	16	7
18										2	9	13	5	18	15	12	14	16

### STEP 3:

Using the original sequence we filled in a diagonal for opponents (from players 1 to 9) for player 10:

round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	11									11	2	9	13	5	18	15	12	14
2		8								8	11	2	9	13	5	18	15	12
3			1							1	8	11	2	9	13	5	18	15
4				3						3	1	8	11	2	9	13	5	18
5					6					6	3	1	8	11	2	9	13	5
6						17				17	6	3	1	8	11	2	9	13
7							4			4	17	6	3	1	8	11	2	9
8								10		10	4	17	6	3	1	8	11	2
9									7	7	10	4	17	6	3	1	8	11
10	16									16	7	10	4	17	6	3	1	8
11		14								14	16	7	10	4	17	6	3	1
12			12							12	14	16	7	10	4	17	6	3
13				15						15	12	14	16	7	10	4	17	6
14					18					18	15	12	14	16	7	10	4	17
15						5				5	18	15	12	14	16	7	10	4
16							13			13	5	18	15	12	14	16	7	10
17								9		9	13	5	18	15	12	14	16	7
18									2	2	9	13	5	18	15	12	14	16

**STEP 4:**

We repeated step 3 for opponents of player 11 - 18:

round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	11	2	9	13	5	18	15	12	14	11	2	9	13	5	18	15	12	14
2	12	8	11	2	9	13	5	18	15	8	11	2	9	13	5	18	15	12
3	18	15	1	8	11	2	9	13	5	1	8	11	2	9	13	5	18	15
4	13	5	18	3	1	8	11	2	9	3	1	8	11	2	9	13	5	18
5	2	9	13	5	6	3	1	8	11	6	3	1	8	11	2	9	13	5
6	8	11	2	9	13	17	6	3	1	17	6	3	1	8	11	2	9	13
7	3	1	8	11	2	9	4	17	6	4	17	6	3	1	8	11	2	9
8	17	6	3	1	8	11	2	10	4	10	4	17	6	3	1	8	11	2
9	10	4	17	6	3	1	8	11	7	7	10	4	17	6	3	1	8	11
10	16	7	10	4	17	6	3	1	8	16	7	10	4	17	6	3	1	8
11	1	14	16	7	10	4	17	6	3	14	16	7	10	4	17	6	3	1
12	6	3	12	14	16	7	10	4	17	12	14	16	7	10	4	17	6	3
13	4	17	6	15	12	14	16	7	10	15	12	14	16	7	10	4	17	6
14	7	10	4	17	18	15	12	14	16	18	15	12	14	16	7	10	4	17
15	14	16	7	10	4	5	18	15	12	5	18	15	12	14	16	7	10	4
16	15	12	14	16	7	10	13	5	18	13	5	18	15	12	14	16	7	10
17	5	18	15	12	14	16	7	9	13	9	13	5	18	15	12	14	16	7
18	9	13	5	18	15	12	14	16	2	2	9	13	5	18	15	12	14	16



## STEP 5:

We assigned the games to the three treatments. Each game is played three times, by three different pairs, in each treatment.

Treatment - A, B or C - is in parentheses.

round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	11(A)	2(A)	9(A)	13(A)	5(A)	18(A)	15(A)	12(A)	14(A)	11(A)	2(A)	9(A)	13(A)	5(A)	18(A)	15(A)	12(A)	14(A)
2	12(B)	8(B)	11(B)	2(B)	9(B)	13(B)	5(B)	18(B)	15(B)	8(B)	11(B)	2(B)	9(B)	13(B)	5(B)	18(B)	15(B)	12(B)
3	18(C)	15(C)	1(C)	8(C)	11(C)	2(C)	9(C)	13(C)	5(C)	1(C)	8(C)	11(C)	2(C)	9(C)	13(C)	5(C)	18(C)	15(C)
4	13(A)	5(A)	18(A)	3(A)	1(A)	8(A)	11(A)	2(A)	9(A)	3(A)	1(A)	8(A)	11(A)	2(A)	9(A)	13(A)	5(A)	18(A)
5	2(B)	9(B)	13(B)	5(B)	6(B)	3(B)	1(B)	8(B)	11(B)	6(B)	3(B)	1(B)	8(B)	11(B)	2(B)	9(B)	13(B)	5(B)
6	8(C)	11(C)	2(C)	9(C)	13(C)	17(C)	6(C)	3(C)	1(C)	17(C)	6(C)	3(C)	1(C)	8(C)	11(C)	2(C)	9(C)	13(C)
7	3(A)	1(A)	8(A)	11(A)	2(A)	9(A)	4(A)	17(A)	6(A)	4(A)	17(A)	6(A)	3(A)	1(A)	8(A)	11(A)	2(A)	9(A)
8	17(B)	6(B)	3(B)	1(B)	8(B)	11(B)	2(B)	10(B)	4(B)	10(B)	4(B)	17(B)	6(B)	3(B)	1(B)	8(B)	11(B)	2(B)
9	10(C)	4(C)	17(C)	6(C)	3(C)	1(C)	8(C)	11(C)	7(C)	7(C)	10(C)	4(C)	17(C)	6(C)	3(C)	1(C)	8(C)	11(C)
10	16(A)	7(A)	10(A)	4(A)	17(A)	6(A)	3(A)	1(A)	8(A)	16(A)	7(A)	10(A)	4(A)	17(A)	6(A)	3(A)	1(A)	8(A)
11	1(B)	14(B)	16(B)	7(B)	10(B)	4(B)	17(B)	6(B)	3(B)	14(B)	16(B)	7(B)	10(B)	4(B)	17(B)	6(B)	3(B)	1(B)
12	6(C)	3(C)	12(C)	14(C)	16(C)	7(C)	10(C)	4(C)	17(C)	12(C)	14(C)	16(C)	7(C)	10(C)	4(C)	17(C)	6(C)	3(C)
13	4(A)	17(A)	6(A)	15(A)	12(A)	14(A)	16(A)	7(A)	10(A)	15(A)	12(A)	14(A)	16(A)	7(A)	10(A)	4(A)	17(A)	6(A)
14	7(B)	10(B)	4(B)	17(B)	18(B)	15(B)	12(B)	14(B)	16(B)	18(B)	15(B)	12(B)	14(B)	16(B)	7(B)	10(B)	4(B)	17(B)
15	14(C)	16(C)	7(C)	10(C)	4(C)	5(C)	18(C)	15(C)	12(C)	5(C)	18(C)	15(C)	12(C)	14(C)	16(C)	7(C)	10(C)	4(C)
16	15(A)	12(A)	14(A)	16(A)	7(A)	10(A)	13(A)	5(A)	18(A)	13(A)	5(A)	18(A)	15(A)	12(A)	14(A)	16(A)	7(A)	10(A)
17	5(B)	18(B)	15(B)	12(B)	14(B)	16(B)	7(B)	9(B)	13(B)	9(B)	13(B)	5(B)	18(B)	15(B)	12(B)	14(B)	16(B)	7(B)
18	9(C)	13(C)	5(C)	18(C)	15(C)	12(C)	14(C)	16(C)	2(C)	2(C)	9(C)	13(C)	5(C)	18(C)	15(C)	12(C)	14(C)	16(C)

## STEP 6:

Subjects' roles were randomly assigned.

\* = P1

round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	11(A)	2(A)	9(A)	13(A)	5(A)	18(A)	15(A)	12(A)	14(A)	11(A)*	2(A)*	9(A)*	13(A)*	5(A)*	18(A)*	15(A)*	12(A)*	14(A)*
2	12(B)	8(B)	11(B)	2(B)	9(B)	13(B)	5(B)	18(B)	15(B)	8(B)*	11(B)*	2(B)*	9(B)*	13(B)*	5(B)*	18(B)*	15(B)*	12(B)*
3	18(C)*	15(C)*	1(C)*	8(C)*	11(C)*	2(C)*	9(C)*	13(C)*	5(C)*	1(C)	8(C)	11(C)	2(C)	9(C)	13(C)	5(C)	18(C)	15(C)
4	13(A)	5(A)	18(A)	3(A)	1(A)	8(A)	11(A)	2(A)	9(A)	3(A)*	1(A)*	8(A)*	11(A)*	2(A)*	9(A)*	13(A)*	5(A)*	18(A)*
5	2(B)	9(B)	13(B)	5(B)	6(B)	3(B)	1(B)	8(B)	11(B)	6(B)*	3(B)*	1(B)*	8(B)*	11(B)*	2(B)*	9(B)*	13(B)*	5(B)*
6	8(C)*	11(C)*	2(C)*	9(C)*	13(C)*	17(C)*	6(C)*	3(C)*	1(C)*	17(C)	6(C)	3(C)	1(C)	8(C)	11(C)	2(C)	9(C)	13(C)
7	3(A)*	1(A)*	8(A)*	11(A)*	2(A)*	9(A)*	4(A)*	17(A)*	6(A)*	4(A)	17(A)	6(A)	3(A)	1(A)	8(A)	11(A)	2(A)	9(A)
8	17(B)*	6(B)*	3(B)*	1(B)*	8(B)*	11(B)*	2(B)*	10(B)*	4(B)*	10(B)	4(B)	17(B)	6(B)	3(B)	1(B)	8(B)	11(B)	2(B)
9	10(C)	4(C)	17(C)	6(C)	3(C)	1(C)	8(C)	11(C)	7(C)	7(C)*	10(C)*	4(C)*	17(C)*	6(C)*	3(C)*	1(C)*	8(C)*	11(C)*
10	16(A)*	7(A)*	10(A)*	4(A)*	17(A)*	6(A)*	3(A)*	1(A)*	8(A)*	16(A)	7(A)	10(A)	4(A)	17(A)	6(A)	3(A)	1(A)	8(A)
11	1(B)*	14(B)*	16(B)*	7(B)*	10(B)*	4(B)*	17(B)*	6(B)*	3(B)*	14(B)	16(B)	7(B)	10(B)	4(B)	17(B)	6(B)	3(B)	1(B)
12	6(C)	3(C)	12(C)	14(C)	16(C)	7(C)	10(C)	4(C)	17(C)	12(C)*	14(C)*	16(C)*	7(C)*	10(C)*	4(C)*	17(C)*	6(C)*	3(C)*
13	4(A)*	17(A)*	6(A)*	15(A)*	12(A)*	14(A)*	16(A)*	7(A)*	10(A)*	15(A)	12(A)	14(A)	16(A)	7(A)	10(A)	4(A)	17(A)	6(A)
14	7(B)	10(B)	4(B)	17(B)	18(B)	15(B)	12(B)	14(B)	16(B)	18(B)*	15(B)*	12(B)*	14(B)*	16(B)*	7(B)*	10(B)*	4(B)*	17(B)*
15	14(C)*	16(C)*	7(C)*	10(C)*	4(C)*	5(C)*	18(C)*	15(C)*	12(C)*	5(C)	18(C)	15(C)	12(C)	14(C)	16(C)	7(C)	10(C)	4(C)
16	15(A)*	12(A)*	14(A)*	16(A)*	7(A)*	10(A)*	13(A)*	5(A)*	18(A)*	13(A)	5(A)	18(A)	15(A)	12(A)	14(A)	16(A)	7(A)	10(A)
17	5(B)	18(B)	15(B)	12(B)	14(B)	16(B)	7(B)	9(B)	13(B)	9(B)*	13(B)*	5(B)*	18(B)*	15(B)*	12(B)*	14(B)*	16(B)*	7(B)*
18	9(C)	13(C)	5(C)	18(C)	15(C)	12(C)	14(C)	16(C)	2(C)	2(C)*	9(C)*	13(C)*	5(C)*	18(C)*	15(C)*	12(C)*	14(C)*	16(C)*

### **Example**

In SESSION 1, subject 1 in round 16 plays game G15, with the display of treatment A, as player P1, and she is paired with subject 13. In round 12 she plays game G6, with the display of treatment C, as player P2, and she is paired with subject 17 (see step 6 table).

In each of the other six sessions we randomized the positions of the rows keeping the distribution of treatments across rows (for example, see the final table for session 2 below).

This matching structure is characterized by the following desirable features:

1. Each game is played 9 times, in 9 different rounds.
2. Each subject plays all 18 games.
3. Each subject plays 2 games against each of 9 different opponents.
4. The order in which games are played is different for each subject.
5. The order in which subject play as P1 or P2 is randomized.
6. Each player plays 6 games in each display.

Matching structure used in SESSION 2 (original position of the row is in brackets)

round	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 [3]	18(A)*	15(A)*	1(A)*	8(A)*	11(A)*	2(A)*	9(A)*	13(A)*	5(A)*	1(A)	8(A)	11(A)	2(A)	9(A)	13(A)	5(A)	18(A)	15(A)
2 [18]	9(B)	13(B)	5(B)	18(B)	15(B)	12(B)	14(B)	16(B)	2(B)	2(B)*	9(B)*	13(B)*	5(B)*	18(B)*	15(B)*	12(B)*	14(B)*	16(B)*
3 [4]	13(C)	5(C)	18(C)	3(C)	1(C)	8(C)	11(C)	2(C)	9(C)	3(C)*	1(C)*	8(C)*	11(C)*	2(C)*	9(C)*	13(C)*	5(C)*	18(C)*
4 [7]	3(A)*	1(A)*	8(A)*	11(A)*	2(A)*	9(A)*	4(A)*	17(A)*	6(A)*	4(A)	17(A)	6(A)	3(A)	1(A)	8(A)	11(A)	2(A)	9(A)
5 [15]	14(B)*	16(B)*	7(B)*	10(B)*	4(B)*	5(B)*	18(B)*	15(B)*	12(B)*	5(B)	18(B)	15(B)	12(B)	14(B)	16(B)	7(B)	10(B)	4(B)
6 [5]	2(C)	9(C)	13(C)	5(C)	6(C)	3(C)	1(C)	8(C)	11(C)	6(C)*	3(C)*	1(C)*	8(C)*	11(C)*	2(C)*	9(C)*	13(C)*	5(C)*
7 [9]	10(A)	4(A)	17(A)	6(A)	3(A)	1(A)	8(A)	11(A)	7(A)	7(A)*	10(A)*	4(A)*	17(A)*	6(A)*	3(A)*	1(A)*	8(A)*	11(A)*
8 [2]	12(B)	8(B)	11(B)	2(B)	9(B)	13(B)	5(B)	18(B)	15(B)	8(B)*	11(B)*	2(B)*	9(B)*	13(B)*	5(B)*	18(B)*	15(B)*	12(B)*
9 [17]	5(C)	18(C)	15(C)	12(C)	14(C)	16(C)	7(C)	9(C)	13(C)	9(C)*	13(C)*	5(C)*	18(C)*	15(C)*	12(C)*	14(C)*	16(C)*	7(C)*
10 [8]	17(A)*	6(A)*	3(A)*	1(A)*	8(A)*	11(A)*	2(A)*	10(A)*	4(A)*	10(A)	4(A)	17(A)	6(A)	3(A)	1(A)	8(A)	11(A)	2(A)
11 [1]	11(B)	2(B)	9(B)	13(B)	5(B)	18(B)	15(B)	12(B)	14(B)	11(B)*	2(B)*	9(B)*	13(B)*	5(B)*	18(B)*	15(B)*	12(B)*	14(B)*
12 [12]	6(C)	3(C)	12(C)	14(C)	16(C)	7(C)	10(C)	4(C)	17(C)	12(C)*	14(C)*	16(C)*	7(C)*	10(C)*	4(C)*	17(C)*	6(C)*	3(C)*
13 [16]	15(A)*	12(A)*	14(A)*	16(A)*	7(A)*	10(A)*	13(A)*	5(A)*	18(A)*	13(A)	5(A)	18(A)	15(A)	12(A)	14(A)	16(A)	7(A)	10(A)
14 [11]	1(B)*	14(B)*	16(B)*	7(B)*	10(B)*	4(B)*	17(B)*	6(B)*	3(B)*	14(B)	16(B)	7(B)	10(B)	4(B)	17(B)	6(B)	3(B)	1(B)
15 [13]	4(C)*	17(C)*	6(C)*	15(C)*	12(C)*	14(C)*	16(C)*	7(C)*	10(C)*	15(C)	12(C)	14(C)	16(C)	7(C)	10(C)	4(C)	17(C)	6(C)
16 [10]	16(A)*	7(A)*	10(A)*	4(A)*	17(A)*	6(A)*	3(A)*	1(A)*	8(A)*	16(A)	7(A)	10(A)	4(A)	17(A)	6(A)	3(A)	1(A)	8(A)
17 [6]	8 (B)*	11(B)*	2(B)*	9(B)*	13(B)*	17(B)*	6(B)*	3(B)*	1(B)*	17(B)	6(B)	3(B)	1(B)	8(B)	11(B)	2(B)	9(B)	13(B)
18 [14]	7(C)	10(C)	4(C)	17(C)	18(C)	15(C)	12(C)	14(C)	16(C)	18(C)*	15(C)*	12(C)*	14(C)*	16(C)*	7(C)*	10(C)*	4(C)*	17(C)*

**Appendix 4: Full results of experiment**

The full results are shown in Table A4.1 below. For each game G1 to G18, there is a panel reporting the frequencies of the choices made by the 126 subjects, disaggregated by player role (P1 or P2) and treatment (A, B or C). Choices are classified according to whether the chosen strategy was displayed in the left (L), right (R) or bottom (B) circle. In each panel, at the top of each ‘treatment’ column, the payoffs of the three pure strategy equilibria are shown in the order L, R, B. In all cases, the cell that reports the frequency of  $s_3$  choices (i.e. the choice predicted by the theory of team reasoning) is shaded grey.

For example, consider game G2, in which the payoffs were  $x = 10, y = 9, v = 8, w = 7$ . In Treatment A, the left, right and bottom circles respectively displayed the  $s_1, s_2$  and  $s_3$  equilibria. This is represented as  $\langle (10, 9), (9, 10), (8, 7) \rangle$ . Of the 21 subjects in the P1 role in this treatment, 10 chose L (i.e.  $s_1$ ), 10 chose R (i.e.  $s_2$ ), and 1 chose B (i.e.  $s_3$ ). Of the 21 subjects in the P2 role in this treatment, 11 chose L (i.e.  $s_1$ ), 10 chose R (i.e.  $s_2$ ), and 0 chose B (i.e.  $s_3$ ). In Treatment B, the left, right and bottom circles respectively displayed the  $s_2, s_3$  and  $s_1$  equilibria. This is represented as  $\langle (9, 10), (8, 7), (10, 9) \rangle$ . Of the 21 subjects in the P1 role in this treatment, 11 chose L (i.e.  $s_1$ ), 0 chose R (i.e.  $s_2$ ), and 10 chose B (i.e.  $s_3$ ). And so on.

At the bottom of each panel, the ‘coordination rate’ is reported for each treatment. This is the number of pairs of co-players in which both players chose the same circle, expressed as a percentage of the 21 pairs.

**Table A4.1: Choice frequencies**

(Shaded cells show  $s_3$  choices.)

GAME G1			
	Treatment A	Treatment B	Treatment C
	$\langle (10,10), (10,10), (9,8) \rangle$	$\langle (10,10), (9,8), (10,10) \rangle$	$\langle (9,8), (10,10), (10,10) \rangle$
L	15	13	3
P1 R	4	4	7
B	2	4	11
TOT	21	21	21

L	10	12	3
P2 R	8	0	10
B	3	9	8
TOT	21	21	21

Coord. Rate (%)			
	38.1	33.3	38.1

GAME G2			
	Treatment A	Treatment B	Treatment C
	$\langle (10,9), (9,10), (8,7) \rangle$	$\langle (9,10), (8,7), (10,9) \rangle$	$\langle (8,7), (10,9), (9,10) \rangle$
L	10	11	0
P1 R	10	0	8
B	1	10	13
TOT	21	21	21

L	11	8	0
P2 R	10	1	11
B	0	12	10
TOT	21	21	21

Coord. Rate (%)			
	38.1	47.6	38.1

GAME G3			
	Treatment A	Treatment B	Treatment C
	$\langle (10,10), (10,10), (9,9) \rangle$	$\langle (10,10), (9,9), (10,10) \rangle$	$\langle (9,9), (10,10), (10,10) \rangle$
L	10	14	2
P1 R	6	2	10
B	5	5	9
TOT	21	21	21

L	11	10	2
P2 R	8	4	14
B	2	7	5
TOT	21	21	21

Coord. Rate (%)			
	52.4	38.1	52.4

GAME G4			
	Treatment A	Treatment B	Treatment C
	$\langle (10,9), (9,10), (8,8) \rangle$	$\langle (9,10), (8,8), (10,9) \rangle$	$\langle (8,8), (10,9), (9,10) \rangle$
L	7	5	5
P1 R	6	13	3
B	8	3	13
TOT	21	21	21

L	5	7	10
P2 R	4	7	3
B	12	7	8
TOT	21	21	21

Coord. Rate (%)			
	28.6	23.8	47.6

GAME G5			
	Treatment A	Treatment B	Treatment C
	$\langle (10,10), (10,10), (10,9) \rangle$	$\langle (10,10), (10,9), (10,10) \rangle$	$\langle (10,9), (10,10), (10,10) \rangle$
L	12	11	4
P1 R	5	1	8
B	4	9	9
TOT	21	21	21

L	15	10	2
P2 R	2	4	10
B	4	7	9
TOT	21	21	21

Coord. Rate (%)			
	47.6	38.1	33.3

GAME G6			
	Treatment A	Treatment B	Treatment C
	$\langle (10,9), (9,10), (9,8) \rangle$	$\langle (9,10), (9,8), (10,9) \rangle$	$\langle (9,8), (10,9), (9,10) \rangle$
L	9	14	3
P1 R	9	0	8
B	3	7	10
TOT	21	21	21

L	15	11	1
P2 R	6	0	11
B	0	10	9
TOT	21	21	21

Coord. Rate (%)			
	52.3	57.1	61.9

GAME G7			
	Treatment A	Treatment B	Treatment C
	<(10,9), (9,10), (9,9)>	<(9,10), (9,9), (10,9)>	<(9,9), (10,9), (9,10)>
L	3	3	12
P1 R	3	16	3
B	15	2	6
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

L	5	3	18
P2 R	5	15	1
B	11	3	2
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

Coord. Rate (%)			
	33.3	66.6	57.1

GAME G8			
	Treatment A	Treatment B	Treatment C
	<(10,10), (10,10), (11,9)>	<(10,10), (11,9), (10,10)>	<(11,9), (10,10), (10,10)>
L	11	12	4
P1 R	4	2	8
B	6	7	9
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

L	15	13	3
P2 R	4	5	10
B	2	3	8
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

Coord. Rate (%)			
	38.1	38.1	52.3

GAME G9			
	Treatment A	Treatment B	Treatment C
	<(12,9), (9,12), (11,10)>	<(9,12), (11,10), (12,9)>	<(11,10), (12,9), (9,12)>
L	4	5	17
P1 R	2	13	1
B	15	3	3
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

L	4	5	15
P2 R	4	14	1
B	13	2	5
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

Coord. Rate (%)			
	57.1	57.1	61.9

GAME G10			
	Treatment A	Treatment B	Treatment C
	<(11,10), (10,11), (12,9)>	<(10,11), (12,9), (11,10)>	<(12,9), (11,10), (10,11)>
L	7	14	3
P1 R	8	1	4
B	6	6	14
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

L	14	5	5
P2 R	2	5	10
B	5	11	6
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

Coord. Rate (%)			
	57.1	23.8	38

GAME G11			
	Treatment A	Treatment B	Treatment C
	<(11,9), (9,11), (10,10)>	<(9,11), (10,10), (11,9)>	<(10,10), (11,9), (9,11)>
L	2	0	19
P1 R	0	19	1
B	19	2	1
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

L	0	3	18
P2 R	1	17	1
B	20	1	2
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

Coord. Rate (%)			
	85.7	71.4	85.7

GAME G12			
	Treatment A	Treatment B	Treatment C
	<(10,10), (10,10), (11,10)>	<(10,10), (11,10), (10,10)>	<(11,10), (10,10), (10,10)>
L	8	9	9
P1 R	1	9	8
B	12	3	4
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

L	3	6	11
P2 R	4	7	8
B	14	8	2
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>

Coord. Rate (%)			
	57.1	19	33.3

GAME G13			
	Treat.A	Treat.B	Treat.C
	<(10,9), (9,10), (11,10)>	<(9,10), (11,10), (10,9)>	<(11,10), (10,9), (9,10)>
L	0	1	16
P1 R	1	18	0
B	20	2	5
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
L	2	3	21
P2 R	1	17	0
B	18	1	0
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
<b>Coord. Rate (%)</b>	80.9	71.4	76.2

GAME G14			
	Treat.A	Treat.B	Treat.C
	<(10,9), (9,10), (10,10)>	<(9,10), (10,10), (10,9)>	<(10,10), (10,9), (9,10)>
L	0	1	20
P1 R	0	20	0
B	21	0	0
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
L	0	1	20
P2 R	0	20	0
B	21	0	0
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
<b>Coord. Rate (%)</b>	100	90.5	90.5

GAME G15			
	Treat.A	Treat.B	Treat.C
	<(10,10), (10,10), (12,11)>	<(10,10), (12,11), (10,10)>	<(12,11), (10,10), (10,10)>
L	4	2	18
P1 R	2	18	2
B	15	1	1
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
L	3	3	11
P2 R	4	14	7
B	14	4	3
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
<b>Coord. Rate (%)</b>	53.2	61.9	61.9

GAME G16			
	Treat.A	Treat.B	Treat.C
	<(10,9), (9,10), (12,11)>	<(9,10), (12,11), (10,9)>	<(12,11), (10,9), (9,10)>
L	1	2	18
P1 R	1	19	2
B	19	0	1
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
L	1	1	18
P2 R	1	20	2
B	19	0	1
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
<b>Coord. Rate (%)</b>	90.5	85.7	85.7

GAME G17			
	Treatment A	Treatment B	Treatment C
	<(10,10), (10,10), (11,11)>	<(10,10), (11,11), (10,10)>	<(11,11), (10,10), (10,10)>
L	1	2	21
P1 R	0	19	0
B	20	0	0
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
L	0	0	20
P2 R	0	20	0
B	21	1	1
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
<b>Coord. Rate (%)</b>	95.2	90.5	95.2

GAME G18			
	Treatment A	Treatment B	Treatment C
	<(10,9), (9,10), (11,11)>	<(9,10), (11,11), (10,9)>	<(11,11), (10,9), (9,10)>
L	0	0	20
P1 R	0	21	0
B	21	0	0
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
L	0	2	20
P2 R	1	19	0
B	20	0	1
<b>TOT</b>	<b>21</b>	<b>21</b>	<b>21</b>
<b>Coord. Rate (%)</b>	95.2	90.5	95.2



### Appendix 5: Correlation between $s_3$ choices in each pair of games

	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12	G13	G14	G15	G16	G17
G1																	
G2	0.1494*																
G3	0.357***	0.136															
G4	-0.027	0.016	0.1677*														
G5	0.529***	-0.054	0.483***	-0.013													
G6	0.125	0.246***	0.310***	0.205**	0.091												
G7	0.034	0.085	0.063	0.451***	0.042	0.087											
G8	0.412***	0.276***	0.369***	-0.025	0.390***	0.253***	0.082										
G9	-0.125	0.085	-0.037	0.070	-0.102	-0.063	0.109	-0.144									
G10	0.002	0.096	0.037	0.004	0.013	0.314***	-0.097	0.190**	-0.140***								
G11	-0.026	0.045	-0.008	0.158*	-0.133	0.086	0.309***	-0.170*	0.418***	-0.204**							
G12	0.177**	0.129	0.2154**	-0.322***	0.250***	0.108	-0.199**	0.341***	0.247	0.028	-0.006						
G13	0.067	0.048	0.011	-0.049	-0.039	0.093	-0.101	0.113	0.054	0.070	0.017	0.28***					
G14	-0.103	0.020	-0.091	-0.073	-0.225**	0.038	0.233***	0.072	0.121	-0.053	0.110	0.153*	-0.060				
G15	0.178**	0.080	0.096	-0.364***	0.217**	0.000	-0.271***	0.198**	0.109	0.182**	-0.112	0.411***	0.233***	-0.099			
G16	0.060	0.047	0.073	-0.220**	0.086	0.089	-0.193**	0.040	0.072	0.121	0.026	0.214**	0.301***	-0.057	0.418***		
G17	0.075	0.026	-0.039	-0.149	0.086	0.049	-0.048	0.094	0.040	-0.001	-0.072	0.119	-0.078	0.235***	0.141	-0.075	
G18	-0.213**	0.023	-0.061	-0.115	-0.050	0.044	0.172*	0.083	0.075	-0.023	0.080	0.178**	-0.069	0.565***	-0.014	0.073	0.19

Pearson correlation index between choices in row game and column game;

\*, \*\* and \*\*\* denote rejection at 10, 5 and 1 significant levels ( $H_0$ : correlation = 0)