Centre for Behavioural and Experimental Social Science

University of East Anglia

2017-02R

GROUP BEHAVIOUR IN TACIT COORDINATION GAMES WITH FOCAL POINTS - AN EXPERIMENTAL INVESTIGATION

Stefania Sitzia

Jiwei Zheng

www.uea.ac.uk/cbess

Group Behaviour in Tacit Coordination Games with Focal Points - An Experimental Investigation*

Stefania Sitzia† Jiwei Zheng‡

January 30, 2018

Abstract

This paper reports an experimental investigation of Schelling's theory of focal points that compares group and individual behaviour. We find that when players' interests are perfectly aligned, groups choose more often the label salient option and achieve higher coordination success than individuals. However, in games with conflict of interest, groups do not always perform better than individuals, especially when the degree of conflict is substantial. We also find that groups outperform individuals in games in which identifying the solution to the coordination problem requires some level of cognitive sophistication (i.e. trade-off games). Finally, players that successfully identify the solution to these games achieve also greater coordination rates in games with a low degree of conflict than other players. This result raises questions of whether finding the focal point is more a matter of logic rather than imagination as instead Schelling argued.

Keywords: Groups; Coordination; Label cues; Cognition

JEL Codes: C72; C78; C91; C92

*This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, grant agreement No. 670103. We thank Robert Sugden, Theodore L. Turocy, Anders Poulsen, and seminar participants at the University of East Anglia for their comments and suggestions, as well participants in the ESA conference in Bergen 2016. All errors are the sole responsibility of the authors.

†Corresponding author. School of Economics and Centre for Behavioural and Experimental Social Science, University of East Anglia, Norwich (UK). Email: s.sitzia@uea.ac.uk.

‡School of Economics and Centre for Behavioural and Experimental Social Science, University of East Anglia, Norwich (UK). Email: j.zheng@uea.ac.uk.

1 Introduction

The assumption of rationality in game theory is used to model behaviour in strategic settings. The theory predicts that the strategy that maximises the agent's utility should be chosen because no other strategy yields a better outcome. It is therefore not surprising that the identity of the player is not taken at all into account in standard game theory; all players behave or should behave rationally. Schelling (1960)'s analysis of focal points is no different in this respect. While his informal experiments are run with individuals, his theory applies to war committees' negotiations as well as to drivers in a traffic jam. It applies to groups as well as individuals. Abundant experimental evidence however shows that groups and individuals do behave differently (e.g. Kugler et al., 2012), so it does matter whether decisions are taken by a board of directors or by a husband trying to meet his wife in New York.

The purpose of this paper is to experimentally investigate to what extent Schelling's theory of focal points applies to groups. We do not claim that our experiment will unveil how war committees agree on limiting wars, or how firms agree on the limits of competition, but we do hope to move a little closer to it. What we will show is that, starting with Schelling's interpretation of coordination games as games that have a solution (e.g. problem-solving tasks), groups have the potential to coordinate more successfully than individuals in tacit coordination games with and without conflict of interest, provided the degree of conflict is small.

Schelling's theory of focal points is supported by many informal experiments involving two-player tacit coordination games. In these games, two players are presented with the same set of n strategies. If they choose the same strategy, they earn some positive payoffs, if they do not, they earn zero. Tacit coordination games can be distinguished into two subsets: games with no conflict of interest and games with conflict of interest.

Games with no conflict of interest, where payoffs are identical in all equilibria and are the same for both players, are pure coordination games. From a game theoretical perspective, the equilibria in these games are indistinguishable to players who play rationally by uniformly randomising across strategies. Schelling reported that in reality people take advantage of some features of the game (irrelevant for game theory and hence called payoff-irrelevant cues, label cues, or conspicuous attributes)[1] and manage to coordinate on a particular equilibrium (the *focal point* of the game) more successfully than if their choices were random. These findings have been corroborated using a variety of games by many experimental studies (e.g. Mehta et al., 1994; Crawford et al., 2008; Isoni et al., 2013; Parravano and Poulsen, 2015).

A distinctive feature of Schelling's theory of focal points lies on the interpretation of coordination games as problems that have a solution. The solution is characterised by being unique and conspicuous (see Lewis, 1969), i.e. the focal point. We can therefore think of these games as problem-solving tasks (e.g. Lorge et al., 1958). These are tasks with a correct solution. In terms of group productivity (Steiner, 1972), tasks can be classified into four types (conjunctive, additive, discretionary, and disjunctive) depending on how individual contributions combine to determine the group's productivity[2]. Problem-solving tasks are *disjunctive tasks* in that the performance of the group is determined by the performance of just one group member. In this type of task, groups are often found to perform better than individuals (e.g. Laughlin et al., 2006). Random samples of two or more individuals are more likely to have an individual able to solve the problem than a random sample of one.

While interpreting pure coordination games as problem-solving tasks seems uncontroversial, doing so for games with conflict of interests requires more words of explanation. Schelling argues that, when conflicts are divergent, parties should just choose following the label cue and reconcile their interest:

> "Beggars cannot be choosers about the source of their signal or about its attractiveness compared with others that they can only wish were as conspicuous. [...] The conflict gets reconciled - or perhaps we should say ignored - as a by-product of the dominant need for coordination." (p.27 and p.22, Schelling, 1957).

[1]Sugden (1995) introduces the concept of labels that are attached to each strategy while (Bacharach, 1993)refers to attributes that characterise an option or object, such as colour or shape.

[2]*Conjunctive tasks* are such that the group productivity is that of the least productive member (e.g. a group climbing a mountain).*Additive tasks* are such that the contributions of the members are added up or averaged to form the group's contribution (e.g. relay race); in *discretionary tasks* group members may combine their effort in any way they like (e.g. musical band).

In this quote, it is clear that the dominant need for coordination is more important than the conflict of interest. If "beggars" want to coordinate, they should just ignore the conflict and follow the "signal". This signal then provides a way to solve the coordination problem and to reconcile the conflict. So, these games, as pure coordination games, are seen by Schelling as games that have a solution.

Contrary to Schelling's expectations, and possibly interpretation, recent experimental evidence (e.g. Crawford et al., 2008; Isoni et al., 2013; Parravano and Poulsen, 2015) has shown that when conflict of interest is introduced, label cues lose much of their power as coordination devices. One possible explanation is that the need for coordination is dominated by the conflict of interest. However, we argue that, when the conflict is negligible, this evidence is not necessarily at odds with Schelling's analysis. In Crawford et al.'s experiment (and Isoni et al.), subjects frequently chose the option that would give them the lower payoff in case of successful coordination (the so-called *after you effect*). We conjecture that subjects were distracted by the payoffs but still wanted to coordinate, even if this meant to choose the option with the lower payoff. We can think therefore of these games as *trade-off* games (Bacharach, 1993). Trade-off games are pure coordination games characterised by the presence of some conspicuous but not unique attribute (or label) and an inconspicuous (or obscure) but unique one[3]. With divergent interests, players' payoffs become more salient than when conflict is absent, and their saliency has the effect of obfuscating that of the label cue. Trade-off games can then be considered as problem-solving tasks with a non-obvious solution. As for pure coordination games, we expect groups to be more successful than individuals at finding the solution in trade-off games and hence in coordination games with small conflict of interests.

Blume and Gneezy (2010) provide experimental evidence using a disc game that possesses the characteristics of a trade-off game. In this game, players are presented with a two-sided 5-sector disc with two black sectors and three white ones. The circular arrangement of the sectors is such

[3]It taken individually conspicuousness and uniqueness, according to Lewis (1969)'notion of saliency, fail to be salient. In fact, an attribute is defined as salient if it is both unique and prominent. If an attribute possesses only one of these two features (i.e. either prominence or uniqueness), then it cannot be considered as such. Bacharach however argued that if we interpret the concept of saliency as one of degrees, then both attributes are salient.

that a white sector is always between the two black ones[4]. The two black sectors are prominent, in that there are only two of them compared to the three white ones while the white sector, in between the two black ones, is unique but not prominent (distinct sector henceforth). Their results show that coordination success in this game is lowered by the difficulty of identification of the distinct sector, but it is higher than that implied by random choices.

Our interpretation of games with a low degree of conflict as trade-off games does not seem applicable when the conflict is large. Crawford et al. (2008) and Isoni et al. (2013) find that in these instances subjects overwhelmingly chose the strategy where they could get, in case of successful coordination, the higher payoff for themselves. This pattern cannot be explained by claiming that it is just the saliency of the label cue that changes. The need for coordination seems to be affected as well. This is not to say that subjects do not want to coordinate, they do want to, but on the equilibrium preferred to them. How groups are expected to perform in these games is not clear. Bornstein and Yaniv (1998) find that in ultimatum games, groups are more rational than individuals in that they propose and accept lower offers; Cox and Hayne (2006) find instead the opposite result in common value auctions. Kocher and Sutter (2005) find that groups are not more rational than individuals in a beauty context game but learn faster. In coordination games with multiple Pareto-ranked equilibria, Feri et al. (2010) find that groups achieve greater coordination on more efficient outcomes and their decisions are more driven by monetary payoffs' concerns. Although evidence on groups' behaviour is mixed, Kugler et al. (2012) conclude that in strategic settings, groups behave more in line with game theoretical assumptions of rationality and selfishness than individuals do. If we were to apply these conclusions to the games we are concerned with, we should expect groups, when the conflict of interest is non-trivial, to coordinate less successfully on the equilibrium suggested by the label cue and choose more often the strategy that, in case of successful coordination, guarantees the higher payoff for themselves.

In this paper, we explore group behaviour using three different games: the pie game (Crawford et al.); bargaining table (Isoni et al.); and the disc game (Blume and Gneezy). The bargaining table and the pie game have been implemented with and without conflict of interest while the disc

[4]Blume and Gneezy (2010) also employ in their experiment another disc game where the two black sectors are next to each other and the three white sectors as well.

game, being a trade-off game, has only be implemented with no conflict of interest to avoid adding any further complexity to the already complex game.

Our main findings are as follows. In pure coordination games, where conflict is absent, groups' choices agree more often with the label cue both when this is conspicuous, as in the pie game and bargaining table, but also when it is not, as in the disc game. When conflict of interest is introduced, groups do better than individuals only in the pie game when the size of the conflict is small. If we restrict our attention to groups and individuals that choose the distinct sector in the disc game, we find that they tend to coordinate more often on the equilibrium suggested by the label cue in both the pie and bargaining games but only when the degree of conflict is small. This suggests that the ability to see the solution when interests are divergent is correlated with the cognitive sophistication needed to notice an inconspicuous but unique cue such as the disc sector.

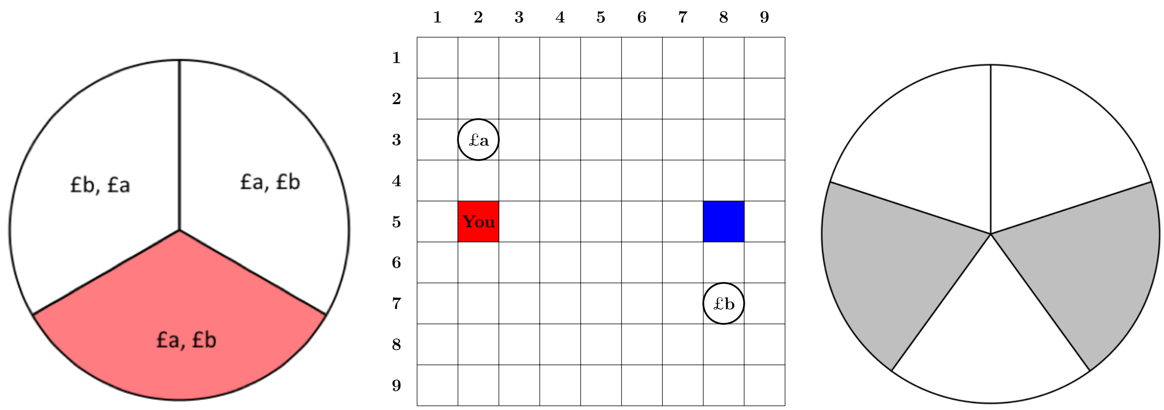
Our paper is organised as follows. Section 2 presents the experimental design; section 3 is devoted to the predictions; sections 4 present the results; and section 5 concludes.

2 Experimental Design

In our experiment, we employed three two-player coordination games: Pie game (Crawford et al.), Bargaining table (Isoni et al.), and Disc game (Blume and Gneezy).

The pie game. In this game, players are presented with a pie (Figure 1a) with three slices of equal size: a red slice, that we will denote as $S1$, and two white slices that we will denote as $S2$ and NS. Players have to choose one of the slices simultaneously and without communication. If they choose the same slice, one earns the amount at the left of the comma while the other earns the amount on the right [5]. The colour red is the label cue of the game. To avoid creating other label cues, such as position, the rotation of the pie is randomised across players.

[5] The payoffs shown on each slice was either $(£a, £b)$ or $(£b, £a)$. From each participant's perspective, the amount of money she could earn was always displayed on the left side of the comma, and the amount the other player could earn was always on the right. In other words, the orders of a and b on the same slice were different between players.



(a) Pie Game (b) Bargaining Table (c) Disc Game

Figure 1: Graphical Representation of the Games Used in the Experiment

We implemented 5 payoffs pairs (Table 1) drawn from a value set $\{a, b\}$, in which $a > b$ (these payoffs represent monetary amounts in pounds). Payoffs $M1$ implement the pie game with no conflict of interest with $a = b$. Payoffs $M2 - M5$ introduce conflict of interest where $a > b$. The payoffs on $S1$ are always the same as the payoffs on the NS while the payoffs on the $S2$ are always reversed.

Payoffs	a	b
M1	5	5
M2	5.1	5
M3	6	5
M4	8	3
M4	10	1

Table 1: Payoffs used in the pie game and bargaining table

The bargaining table. In the bargaining table, two players are presented with a 9×9 square grid with two coloured bases (red and blue) and two discs. Players are randomly assigned to one of the bases, and they select simultaneously and without communication only one of the two discs (we will say they claim a disc). If they claim the same disc, they earn nothing. If they claim different discs, they earn a positive payoff shown on the disc itself (see Figure 1b for a graphical representation of the bargaining table). The bases were always placed in row 5 and column 2 and 8, respectively (column and row numbers were not shown to subjects). The colour of the bases was randomised, so that the blue base was sometimes on the left of the table and sometimes on the right. The two discs were always placed in row 3, column 2 and in row 7, column 8, respectively, so that each disc was relatively close to one of the bases. In bargaining tables, closeness is the label cue that has been shown to be a powerful coordination device (e.g. Isoni et al., 2013) The payoffs we employed in this game are the same as the ones employed in the pie game.

The disc game. The disc game is the only trade-off game in our experiment. In this game, a pair of players is presented with a two-sided disc with two black sectors and three white ones. The graphical representation of the disc is shown in Figure 1c. The side and the rotation of the disc are randomised across players. Players have to choose simultaneously and without communication one of the sectors. If they both choose the same sector, they earn a positive amount; otherwise, they earn nothing. With such a complex game, we decide to implement only payoffs $M1$ and not to display the monetary amounts on the sectors, as this was deemed not necessary. Given that subjects could be presented with either side of the disc and any rotation, the only sector that could be uniquely identified was the white sector between the two black ones (henceforth distinct sector).

Our experiment involves two treatments, an individual treatment (**I**) and a group treatment (**G**) implemented between subjects. The main difference between the two treatments is that in the **I** treatment players are individuals, while in the **G** treatment players are groups of two individuals.

*The **I** treatment.* The **I** treatment is our baseline treatment. Each subject was randomly and anonymously paired with another subject in the room and played the 11 coordination games in random order. No communication was allowed during the experiment, and no feedback was given until the end.

*The **G** treatment.* In the **G** treatment players are groups of two individuals. At the beginning of each session, every participant was informed that she would be randomly paired with another participant in the room to form a group, and that this person would be the same for the whole duration of the experiment. It was common knowledge that players were groups of two individuals. A group-decision was reached only when both members agreed on the same option. The process of reaching a group decision was the following. First, each group member privately suggested what the group should choose. Once both members gave their suggestions, they could see the options they had suggested on their screen. At this stage, they were allowed to communicate via a chat box to concert their choices. Once they agreed unanimously on which option to choose, each group member selected on their respective screen the option they agreed on; a second screen would then show a message confirming that their choices coincided, and the group could then proceed to the next game. If for any reason, choices did not coincide, the group had the chance to chat again to revise their choices. This process could be repeated 5 times. If no agreement was reached by then, the computer would randomly choose on their behalf[6]. To keep monetary incentives the same between treatments, earnings per participant in the **G** treatment were the same as the corresponding earnings in the **I** treatment. For example, if two groups coordinated successfully on a game with no conflict of interest, each group member would earn £5. We ran the experiment at the University of East Anglia in February 2016. We recruited 148 subjects (48 subjects in the **I** treatment and 50 groups of two individuals each in the **G** treatment) with the online recruiting system hRoot (Bock et al., 2014). Subjects were asked to read the instructions and to answer a brief questionnaire to make sure that instructions were correctly understood. Once all clarification questions were answered, the experiment started. At the end of the experiment, one of the games was randomly chosen, and subjects were paid the amount they had earned in that game. Additionally, they were given a participation fee of £5.

3 Predictions

In this section, we will derive some hypotheses that will be used as a benchmark to compare behaviour of groups versus the behaviour of individuals. We will use the theory of team reasoning to draw hypotheses for the pie game and the bargaining table, and Variable Frame Theory (VTF, Bacharach, 1993; Bacharach and Bernasconi, 1997) for the trade-off game (i.e. disc game). Both theories are consistent with Schelling's view of coordination games as problems with a solution hence they seem to be the most natural choices. Although we could employ team reasoning to derive predictions in the disc game, VFT is the best-suited theory, as it has been developed by Bacharach to explain behaviour in tacit coordination games including those where cognition might play a role.

[6]In the actual experiment, all groups reached a final decision. In fact, only few groups needed more than 1 chat sessions to reach a common decision. In the few cases where two chat sessions were needed, it was because one the group members changed her mind at the last minute.

Cognitive hierarchic (e.g. Camerer et al., 2004) and Level-K models (Crawford et al., 2008) are also often used to explain behaviour in coordination games. These models assume that players differ in their level of cognitive sophistication. Level-K theory, for example, assumes Level-0 players choose non-strategically, and higher levels best respond to players that are just one level below them. Although these models can be used to derive predictions in the games we employ in this experiment, assumptions regarding the distribution of levels in the population, as well as how levels combine to form a group level, need to be introduced. The predictions so derived are however sensitive to these auxiliary assumptions, hence we will just employ team reasoning and VFT.[7]

3.1 Pie Game and Bargaining Table

Although some versions of team reasoning are different in respect of whether individuals aim to maximise team's utility (Bacharach, 2006; Sugden, 1993, 1995) or achieve mutual benefits (Sugden, 2018), their core is very similar. The theory of team reasoning posits that individuals will make choices based on team thinking, that is, instead of asking 'what should I do?' they ask 'what should we do?'. They first try to find a rule (*the best rule* - e.g. a strategy profile in a game), if exists, that either maximises the team utility or yields mutual benefits. Each individual then follows that rule in the expectation that others will do the same. In 2-player pure coordination games, such a rule is for both players to choose the strategy suggested by the label cue. When conflict of interest is introduced, team-reasoning predictions do not change.

Team reasoning theories were developed as an attempt to formalise Schelling's theory of focal point, and as Schelling's theories they do seem to consider coordination games as problems that have a correct solution (i.e. the best rule), that is, problems-solving tasks. Evidence in social psychology shows that groups do better than individuals in this type of tasks (Laughlin et al., 2006) hence we expect groups to be more likely than individuals to follow the best rule as they are more likely to find it.

***H1**: When interests are aligned, in both pie game and bargaining table, team reasoning predicts that groups will choose more often than individuals the label salient option.*

[7] See Appendix B for more details about the limitations of Level-K predictions.

H2*: When interests are divergent, in both pie game and bargaining table, team reasoning predicts that groups will choose the label salient option more often than individuals.*

3.2 Disc game

A central concept of VFT is that of frame. In coordination games, players are usually required to choose one out of a set of n objects. These objects are characterised by some attributes. A frame is defined as a set of families of attributes which can be used to partition the objects or options into subsets. In the disc game, we can identify two families of attributes: the colour family (C) and the circular order of the sectors (O). There are therefore four possible frames associated with these families of attributes: the empty frame $\{\varnothing\}$, the colour frame $\{C\}$, the circular order frame $\{O\}$, and the conjunction of the last two frames $\{C, O\}$. Only the last frame is complete while the others are subsets of this.

In VFT, frames determine which options players have available to choose from. If a player has only the colour frame $\{C\}$, her options are {choose a sector, choose a black sector, choose a white sector}. If a player has the complete frame $\{C, O\}$ she is able to notice, given the curricular arrangement of sectors and their colour, that one of them is unique (the distinct white sector), her options are then {choose a sector, choose a black sector, choose a white sector, choose the distinct sector}.

Attributes can be distinguished depending on their availability. When attributes comes easily to mind, they are said to be easily available. We define as v the probability that a player notices (is aware of) an attribute. In the disc game, colour is an easily available attribute, so it seems uncontroversial to assume that all players have the colour frame $\{C\}$, hence $v = 1$. The fact, instead, that the circular order of the sectors in conjunction with the colour implies that the white sector in between the black one is a unique one, is not easily available, hence $v < 1$. The probability v, in VFT, is independent across players. There will therefore be players that have a complete frame $\{C, O\}$ and players that only have a subset of it, that is, $\{C\}$. Players' beliefs about other players' frames are correct but limited by their own frame, that is, a player who only has the $\{C\}$ frame cannot have a belief that other players have the complete frame $\{C, O\}$, because she does not have that frame. Given a frame, players are assumed to choose the option whose attribute is rarer because this guarantees higher chances of successful coordination (Rarity Preference). In the disc game, if a player has the $\{C\}$ frame, she will pick one of the black sectors, as there are only two of them, while there are three white ones. If a player has a complete frame $\{C, O\}$ she will choose the distinct sector only if she believes that the probability v that the other player has a complete frame is greater than the probability of coordinating on the black sectors. Given that there are only two black sectors, a player with a complete frame will choose the distinct sector only if $v > 1/2$.

We will now adapt this framework to obtain predictions for the G treatment. When players are randomly paired in groups, the probability v that the distinct sector will be noticed increases as only one player in the group is needed for the group to notice it and v is independent across players. In addition, given that group formation is common knowledge, groups' beliefs (that are assumed to be correct albeit limited by their frame) about the probability that other groups will notice the distinct sector will increase compared to individuals. Both factors lead to our third hypothesis.

H3*: VFT predicts that, in the disc game, groups will choose more often than individuals the distinct sector.*

4 Results

4.1 Overview

Figure 2 reports the proportion of times the label cue is chosen by treatment, game type and payoff type. In all games and for all payoffs, except $M5$ for the bargaining table (BT), groups choose consistently with the label cue more often than individuals. In both treatments, these proportions decrease as the difference in payoffs increases ($M1 - M5$). Similarity the difference between treatments becomes less sharp.

Table 2 presents the same summary statistics as Figure 2 but with greater level of detail. We will refer to this table repeatedly in this section. The table is split into three panels. **Panel a**

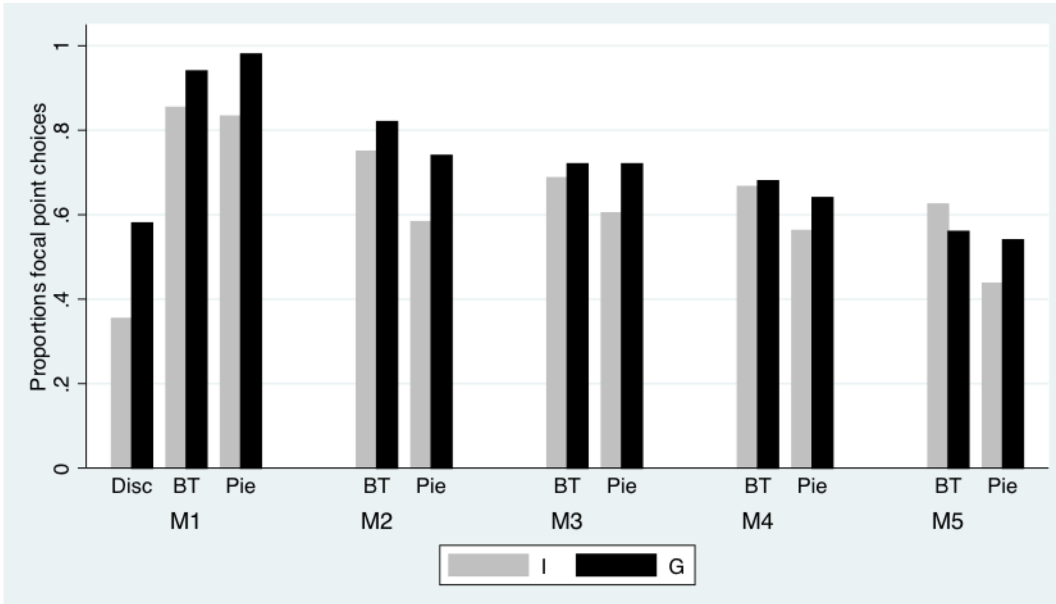


Figure 2: Proportion of focal point choices by treatment, game and payoff type.

reports summary results for the pie game, **Panel b** for the bargaining table and **Panel c** for the disc game. For each game, we report the distribution of choices over options (slices, discs, and sectors). The slices in the pie game are labelled $S1$ (the red slice with payoffs (a, b)), $S2$ (the white slice with payoffs (b, a)) and NS (the white slice with payoffs (a, b) - the same as $S1$).

The sectors in the disc game are labelled as D for the distinct sector, $B1$ and $B2$ for the two black sectors and $W1$ and $W2$ for the two white adjacent sectors. For the pie game and the bargaining table, we report choices broken down by payoff type and also player type for games with conflict of interests (payoffs $M2 - M5$). We define as Player 1 ($P1$) the player whose payoff is greater in the option suggested by the label cue and Player 2 ($P2$) the player whose payoff is instead lower. Finally, for each payoff pair we report the expected coordination rates (ECR), and the significance level of their difference between treatments. The ECRs are obtained using the following formula (Mehta et al., 1994):

$$ECR = \sum_i ECR_i = \sum_i \frac{n_i (n_i - 1)}{N (N - 1)}$$

where N is the number of players (individuals or groups, depending on the treatment), i is the slice, disc, or sector we are calculating the ECR of, and n_i is the number of players that have chosen

Panel a: Pie Game

Treatment	Slice	Player	M1	M2	M3	M4	M5
			(5, 5)	(5.1, 5)	(6, 5)	(8, 3)	(10, 1)
I	S1	P1		0.50	0.63	0.63	0.54
		P2		0.67	0.58	0.50	0.33
		Pooled	*0.83*	*0.58*	*0.60*	*0.56*	*0.44*
	S2	P1		0.46	0.33	0.29	0.21
		P2		0.04	0.21	0.33	0.67
		Pooled	*0.06*	*0.25*	*0.27*	*0.31*	0.44
	NS	P1		0.04	0.04	0.08	0.25
		P2		0.29	0.21	0.17	0.00
		Pooled	*0.10*	*0.17*	*0.13*	*0.13*	0.13
G	S1	P1		0.72	0.80	0.84	0.80
		P2		0.76	0.64	0.44	0.28
		Pooled	*0.98*	*0.74*	*0.72*	*0.64*	*0.54*
	S2	P1		0.24	0.08	0.16	0.12
		P2		0.24	0.36	0.56	0.64
		Pooled	*0.00*	*0.24*	*0.22*	*0.36*	*0.38*
	NS	P1		0.04	0.12	0.00	0.08
		P2		0.00	0.00	0.00	0.08
		Pooled	*0.02*	*0.02*	*0.06*	*0.00*	*0.08*
		Expected Coordination Rates					
I			0.70	0.36	0.44	0.42	0.32
G			0.96	0.60	0.54	0.46	0.31
		SL	***	***	*	ns	ns

Panel b: Bargaining Table

Treatment	Player	M1	M2	M3	M4	M5
		(5, 5)	(5.1, 5)	(6, 5)	(8, 3)	(10, 1)
I	P1		0.75	0.58	0.58	0.79
	P2		0.75	0.79	0.75	0.46
	Pooled	*0.85*	*0.75*	*0.69*	*0.67*	*0.63*
G	P1		0.72	0.72	0.76	0.88
	P2		0.92	0.72	0.60	0.24
	Pooled	*0.94*	*0.82*	*0.72*	*0.68*	*0.56*
	Expected Coordination Rates					
I		0.73	0.63	0.55	0.54	0.48
G		0.88	0.68	0.60	0.55	0.30
	SL	**	ns	ns	ns	***

Panel c: Disc Game

Treatment	D	B1	B2	W1	W2	*Total*
I	0.35	0.21	0.31	0.08	0.04	*1.00*
G	0.58	0.22	0.16	0.02	0.02	*1.00*
	Expected Coordination Rates					
I	0.12	0.04	0.09	0.01	0.00	*0.26*
G	0.33	0.04	0.02	0.00	0.00	*0.40*
SL	***	ns	ns	ns	ns	***

Table 2: Frequency of choices in all games by treatment and payoff type
Significance levels: ns. not significant, * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

that slice, disc, or sector. The ECR_i are then added up to obtain at total ECR for a particular game and payoff pair. Given that each ECR consists of only one observation, to test whether there is a treatment effect, we use the bootstrap procedure employed by Bardsley et al. (2010). We repeatedly sample with replacement 10,000 times the distribution of choices for each payoff pair and game type in the **I** treatment. For each of these distributions, we calculate the corresponding ECR as explained above. We then use this distribution to obtain confidence intervals for the **G** treatment.

4.2 The pie game and the bargaining table.

Games with no conflict of interests. In games with no conflict of interests (payoff $M1$) groups' choices agree with the label cue more often than individuals' ones. In the pie game (**Panel a**), 98% of groups choose the red slice compared to 83% of individuals (test of proportions, $p < 0.01$); in the bargaining table (**Panel b**), 94% of the groups choose the close disc compared to 85% of individuals (test of proportions, $p < 0.09$). Binomial test results reveal that these proportions are significantly greater than those entailed by random choices ($p < 0.01$ in both cases). The ECRs are shown at the bottom of each panel. They are extremely high in both treatments but exceptionally high in the **G** treatment (96% in the pie game and 88% in the bargaining table). The difference in ECRs between treatments is significant (bootstrap procedure, $p < 0.001$ and $p < 0.05$ for the pie game and bargaining table, respectively). This result is consistent with *H1*.

R1*: In the pie game and bargaining table, groups coordinate more often than individuals on the focal point equilibrium when there is no conflict of interest.*

Games with conflict of interest. When conflict of interest is introduced, the pattern in the data is more complex. In the pie game, pooling across $P1$ and $P2$, we observe that the $S1$ slice is chosen less frequently in both treatments as the payoff difference increases. The most notable difference between treatments is that groups choose significantly less often than individuals the NS slice (Chi2 test, $p < 0.001$) and as a consequence, they concentrate more their choices on the two remaining slices $S1$ and $S2$. This result seems to suggest that groups do indeed perceive focality more strongly than individuals, this allows them to exclude from their choices the slice NS and to achieve greater ECRs than individuals for payoffs $M2$ and $M3$ (see Table 2). When the payoff difference increases, ECRs not surprisingly drop in both treatments.

In the bargaining table, both groups and individuals choose less often according to the label cue as the payoff difference increases. Groups tend to choose more often than individuals the close disc (except for payoff $M5$) but the difference is small and not significant for any payoff pair we employ. ECRs decrease as payoff difference increases and, unlike in the pie game, there are no differences across treatments (the only exception is $M5$ where groups do worse than individuals). Given that in this game there are only two options to choose from (unlike in the pie game where there are three options), groups' stronger sensitivity to focality observed in the pie game, is of no help. This result does not provide full support to $H2$.

R2*: Although groups seem to perceive saliency more strongly than individuals, this is not enough to fully support H2.*

4.3 Disc Game

The disc game is a trade-off game (Bacharach and Bernasconi, 1997) in that the identification of the unique distinct sector requires some level of cognitive sophistication. In line with $H3$, groups chose the distinct sector (58%) more often than individuals (35%). This difference is significant at a 5% level (test of proportions). Individuals (52%) chose more often the black sectors ($B1 + B2$) than groups while the two adjacent white sectors ($W1 + W2$) were chosen less frequently by both groups and individuals. In the bootstrapped results for the D sector, only 8 observations out of 10,000 ($p < 0.001$) are greater than the observed ECR_D in the **G** treatment. As a consequence, the difference in ECR is as well strongly significant (see Table 2, **Panel c**).

R3a*: In line with H3, groups choose significantly more often than individuals the distinct sector in the disc game, and this leads to a greater ECR than that in the I treatment.*

To understand what is driving this result, we have looked at the decisions each group member made before having the chance to chat with their own partner. Before chatting, the proportion of group members that chose the distinct sector was 42%, a bit greater than that in the **I** treatment but not significantly different (test of proportion, $p = 0.85$). This shows that at the individual level, choices in the two treatments are similar. In only 14 cases, group members chose the same sector before chatting and in 11 of these cases they chose the distinct sector. All these groups confirmed their initial choice after chatting. Of the 36 groups whose members disagreed on their initial choices, 20 had one member who suggested the distinct sector before communication. In 17 of these cases, the group finally agreed on choosing the distinct sector, and only in 3 cases the group opted for one of the black sectors. From the analysis of the chats we found evidence that 12 of these 17 groups chose the distinct sector because of its uniqueness. In the other 5 cases, the analysis of the chats is inconclusive. Finally, the groups that disagreed on sectors other than the distinct one reached a common decision because one of the members changed his/her suggestion to match the choice of the other member. In these cases, the chat analysis is inconclusive. But this is not surprising, as both black sectors and the two adjacent white ones are not unique, no argument can be provided for favouring one over the other. Just in few cases groups referred to the black sectors as being the right choice, or that coordination in that game was just random. This result provides support to the interpretation of coordination games as problems with a solution. If one of the member notices the uniqueness of the distinct sector, the other member is persuaded that choosing that sector is the *correct* choice. The chat analysis also shows that beliefs, about whether the other player has noticed the distinct sector, do not seem to play a crucial role, once one of the group members finds the solution the group will adopt it as such.

R3b*: If one of the group members is aware of the distinct sector, the majority of groups chooses the distinct sector.*

4.4 Analysis by type

In this section, we ask whether the cognitive sophistication needed for identifying the distinct sector in the disc game is correlated with choices that agree with the label cue in the pie game and bargaining table. The intuition is that groups and individuals that choose the distinct sector (D-type) show to possess the necessary cognitive sophistication to solve coordination problems that might therefore be transferred to other contexts such as the pie game and bargaining table. Cognitive sophistication can be thought of as a trait that individuals or groups possess and is observed in the disc game that acts as a measuring device.

It might be argued that the Pie game and the bargaining table feature an obvious label cue hence cognition does not really play a role. However, we have contended in the introduction that, despite its saliency, the solution when there is conflict of interest, in particular when this is negligible (i.e. $M2$ and $M3$), is obfuscated by the payoff difference, hence relatively less salient. These games can therefore be interpreted as trade-off games such as the disc game. When the payoff difference is large ($M4$ and $M5$), both groups and individuals still want to coordinate but on their preferred outcome rather than the one suggested by the label cue. This seems to be supported by the fact (Table 3) that an overwhelming proportion of both groups and individuals choose the option with the higher payoff for themselves. Hence these games cannot be interpreted as trade-off games.

We define as D-type groups or individuals who have chosen the distinct sector (46 observations) and O-types those who have not (52 observations). Figure 3 and Table 3 present the frequency of choices in the pie game and bargaining table that agree with the label cue, broken down by treatment, payoff pair, and type. At the aggregate level, we found strong evidence that D-types choose according to the label cue more often than O-types (Mann-Whitney, $p < 0.02$).

Mann-Whitney test results also reveal that this proportion is significantly greater for D-groups than D-individuals and O-type both individuals and groups ($p < 0.09$, $p < 0.001$ and $p < 0.04$, respectively). Pairwise comparisons between D-individuals and O-types are instead not significantly different.

R4: *D-type groups choose the label salient option significantly more often than D-type individuals, O-type groups, and O-type individuals.*

Panel a: frequency of focal point choices by type

Treatment	player	type	Bargaining Game M1	M2	M3	M4	M5	Pie Game M1	M2	M3	M4	M5
			(5, 5)	(5.1, 5)	(6, 5)	(8, 3)	(10, 1)	(5, 5)	(5.1, 5)	(6, 5)	(8, 3)	(10, 1)
I	P1	O		0.75	0.53	0.60	0.93		0.40	0.61	0.61	0.59
		D		0.75	0.67	0.56	0.56		0.67	0.67	0.67	0.43
	P2	O		0.67	0.75	0.69	0.50		0.69	0.62	0.38	0.43
		D		0.89	0.88	0.88	0.38		0.63	0.55	0.64	0.20
	pooled	*O*	*0.84*	*0.71*	*0.65*	*0.65*	*0.71*	*0.81*	*0.55*	*0.61*	*0.52*	*0.52*
		D	*0.88*	*0.82*	*0.76*	*0.71*	*0.47*	*0.88*	*0.65*	*0.59*	*0.65*	*0.29*
G	P1	O		0.54	0.58	0.73	0.85		0.50	0.70	0.67	0.75
		D		0.92	0.85	0.80	0.92		0.87	0.87	0.94	0.85
	P2	O		1.00	0.56	0.50	0.13		0.64	0.55	0.33	0.22
		D		0.88	0.81	0.63	0.29		0.86	0.71	0.54	0.31
	pooled	*O*	*0.95*	*0.71*	*0.57*	*0.67*	*0.57*	*0.95*	*0.57*	*0.62*	*0.48*	*0.52*
		D	*0.93*	*0.90*	*0.83*	*0.69*	*0.55*	*1.00*	*0.86*	*0.79*	*0.76*	*0.55*

Panel b: ECRs by type

Treatment	Cognition	Bargaining Game M1	M2	M3	M4	M5	Pie Game M1	M2	M3	M4	M5
		(5, 5)	(5.1, 5)	(6, 5)	(8, 3)	(10, 1)	(5, 5)	(5.1, 5)	(6, 5)	(8, 3)	(10, 1)
I	O	0.73	0.58	0.52	0.54	0.50	0.69	0.28	0.44	0.35	0.32
	D	0.79	0.69	0.63	0.54	0.49	0.79	0.47	0.45	0.55	0.43
	p-values		0.09	0.06	0.41	0.54		0.04	0.42	0.01	0.10
	Significance		*	*	ns	ns		**	ns	***	*
G	O	0.91	0.54	0.51	0.50	0.24	0.91	0.46	0.43	0.44	0.29
	D	0.87	0.82	0.72	0.58	0.33	1.00	0.76	0.64	0.53	0.32
	p-values		0.02	0.01	0.21	0.24		0.00	0.05	0.08	0.40
	Significance		***	***	ns	ns		***	**	*	ns

Table 3: Frequency of focal point choices and ECRs by game and payoff type. Significance levels: ns. not significant, * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

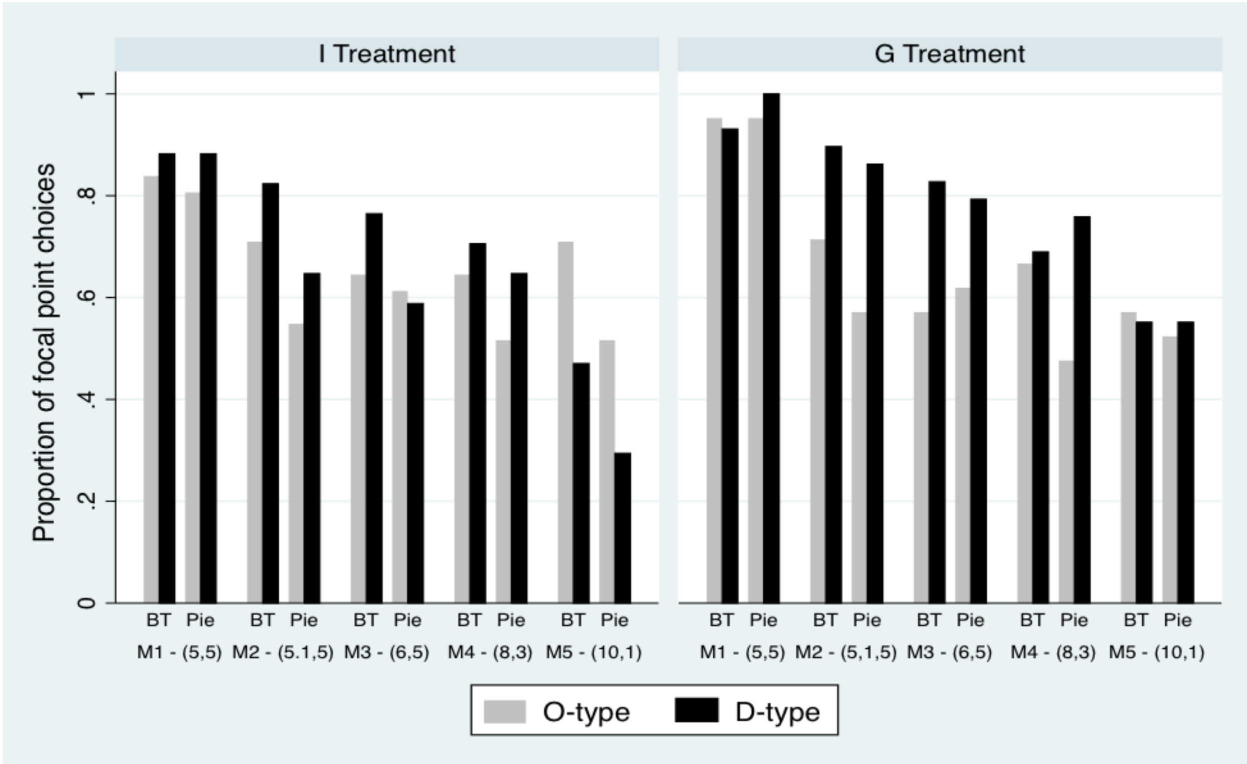


Figure 3: Proportion of focal point choices by type.

Panel b of Table 3 reports the ECRs per treatment and payoff pair obtained by matching D-players with D-players and O-players with O-players. D-type's ECRs are in most cases greater than O-type's ones. In the **I** treatment the difference in ECRs between types is less pronounced than in the **G** treatment but it is almost always significant. In the **G** treatment, the difference in ECRs between types is significant except for payoffs $M4$ and $M5$. Interestingly, these ECRs do not significantly differ from the ECRs obtained in the **I** treatment for games with no conflict of interest. This suggests that, for D-groups, small payoff differences are not an issue and do not lead to coordination failure.

The results in this section lend some support to the hypothesis that games with small payoff differences can be thought of trade-off games. In these games, D-groups, and to a lesser extent D-individuals, are not distracted by small payoff differences and additionally have the necessary cognitive sophistication to identify the unique solution. Our hunch of why cognition might play a role lies in the fact that D-types might be particularly sensitive to symmetry. That is, to see the distinct sector, they should notice that there is an asymmetry and this lies in the distinct sector.

In the pie game and bargaining table, Player 1 and Player 2 are in exactly the same position but with opposite interests. A person sensitive to symmetry might be able to see, form a purely logical point of view, that whatever applies to Player 1 applies, but in opposite direction, to Player 2. Hence the unique rule that can accommodate this difference is to choose a rule that is not based on payoffs and is the same for both players (i.e. symmetric). This rule is the one dictated by the label cue.

D-groups are more successful in coordination than D-individuals. We reckon that a possible reason might relate to the way we measure cognition. Firstly, some participants might have chosen the distinct sector randomly (this might account only for a small number of choices). In the **G** treatment, however, we have provided evidence that most groups chose the distinct sector because of its uniqueness rather than randomly. Secondly, VFT predicts that the belief that other players have noticed the distinct sector is smaller in the **I** treatment than in the **G** treatment. In the **G** treatment, from the chat analysis, evidence suggests that beliefs did not play a substantial role in decisions, however we do not have this evidence in the **I** treatment. Hence it might be possible that the number of participants choosing a black sector, even if aware of the distinct one, is greater in the **I** treatment than in the **G** treatment. Both these factors together make the D-type/O-type distinction in the **I** treatment less precise than in the **G** treatment.

5 Conclusions

In this paper, we asked whether Schelling's theory of focal points can be extended to groups. We have argued that, consistently with Schelling's view, coordination games with label cues can be thought of as games with a correct solution and this should lead groups to coordinate more successfully than individuals.

In line with our predictions, we find that groups outperform individuals when interests are aligned and when some level of cognitive sophistication is required to identify the solution to the problem of coordination.

When interests are not aligned, and consistently with previous findings (e.g. Crawford et al., 2008; and Isoni et al., 2013), the expected coordination rates in both treatments drop. Our analysis though suggests that groups seem to perceive the label cue more strongly than individuals, as evidenced by the results in the pie game. This phenomenon is not detectable in the bargaining table where the choice is only between two strategies, close disc and far disc.

Finally, groups and individuals who choose more frequently the distinct sector follow the label cue more often than those who have chosen any other sector. We advanced a conjecture based on cognition. Cognition endows players with the ability to identify the perfectly anti-symmetrical position of their co-player compared to theirs and find the rule that applies equally to both of them. This does not necessarily imply maximising the team's utility function but stems mainly from logical considerations arising from the structure of the game. For these groups and individuals, uniqueness might be so salient to overshadow that saliency of the payoffs.

Schelling (1960) believed that finding the key to solve a coordination problem involved more imagination than logic, however we have shown that this is only part the story. Groups and cognitive sophisticated players display a better ability in finding the solution to the riddle of coordination. This implies that coordination success is not just a matter of wanting to coordinate but also being able to do so, and groups have shown in this respect a greater ability.

References

Michael Bacharach. Variable universe games. In Ken Binmore, Alan Kinmar, and Piero Tani, editors, *Frontiers of game theory*, pages 255–276. Mit Press, Cambridge, 1993.

Michael Bacharach. *Beyond individual choice: teams and frames in game theory*. Princeton University Press, 2006.

Michael Bacharach and Michele Bernasconi. The variable frame theory of focal points: An experimental study. *Games and Economic Behavior*, 19(1):1–45, 1997.

Nicholas Bardsley, Judith Mehta, Chris Starmer, and Robert Sugden. Explaining Focal Points: Cognitive Hierarchy Theory versus Team Reasoning. *The Economic Journal*, 120(543):40–79, 2010.

Andreas Blume and Uri Gneezy. Cognitive forward induction and coordination without common knowledge: An experimental study. *Games and Economic Behavior*, 68(2):488–511, 2010.

Olaf Bock, Ingmar Baetge, and Andreas Nicklisch. hroot: Hamburg registration and organization online tool. *European Economic Review*, 71:117–120, 2014.

Gary Bornstein and Ilan Yaniv. Individual and group behavior in the ultimatum game: Are groups more "rational" players? *Experimental Economics*, 1(1):101–108, 1998.

Colin F Camerer, Teck-Hua Ho, and Juin-Kuan Chong. A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3):861–898, 2004.

Miguel A. Costa-Gomes and Vincent P. Crawford. Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review*, 96(5):1737–1768, 2006.

James C Cox and Stephen C Hayne. Barking up the right tree: Are small groups rational agents? *Experimental Economics*, 9(3):209–222, 2006.

Vincent P Crawford, Uri Gneezy, and Yuval Rottenstreich. The power of focal points is limited: even minute payoff asymmetry may yield large coordination failures. *The American Economic Review*, 98(4):1443–1458, 2008.

Francesco Feri, Bernd Irlenbusch, and Matthias Sutter. Efficiency gains from team-based coordinationlarge-scale experimental evidence. *American Economic Review*, 100(4):1892–1912, 2010.

Andrea Isoni, Anders Poulsen, Robert Sugden, and Kei Tsutsui. Focal points in tacit bargaining problems: Experimental evidence. *European Economic Review*, 59:167–188, 2013.

Martin G Kocher and Matthias Sutter. The Decision Maker Matters: Individual Versus Group Behaviour in Experimental Beauty-Contest Games*. *The Economic Journal*, 115(500):200–223, 2005.

Tamar Kugler, Edgar E Kausel, and Martin G Kocher. Are groups more rational than individuals? A review of interactive decision making in groups. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(4):471–482, 2012.

Patrick R Laughlin, Erin C Hatch, Jonathan S Silver, and Lee Boh. Groups perform better than the best individuals on letters-to-numbers problems: effects of group size. *Journal of Personality and social Psychology*, 90(4):644–651, 2006.

David Lewis. Convention cambridge. *Mass.: Harvard UP*, 1969.

Irving Lorge, David Fox, Joel Davitz, and Marlin Brenner. A survey of studies contrasting the quality of group performance and individual performance. *Psychological bulletin*, 55(6):337–372, 1958.

Judith Mehta, Chris Starmer, and Robert Sugden. The nature of salience: An experimental investigation of pure coordination games. *The American Economic Review*, 84(3):658–673, 1994.

Melanie Parravano and Odile Poulsen. Stake size and the power of focal points in coordination games: Experimental evidence. *Games and Economic Behavior*, 94:191–199, 2015.

Stefan P Penczynski. Persuasion: An experimental study of team decision making. *Journal of Economic Psychology*, 56:244–261, 2016.

Thomas C Schelling. Bargaining, communication, and limited war. *Conflict Resolution*, 1(1):19–36, 1957.

Thomas C Schelling. The strategy of conflict. *Cambridge, Mass*, 1960.

Dale O Stahl and Paul W Wilson. On players models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1):218–254, 1995.

ID Steiner. Group processes and group productivity. *New York: Academic*, 1972.

Robert Sugden. Thinking as a team: Towards an explanation of nonselfish behavior. *Social Philosophy and Policy*, 10(1):69–89, 1993.

Robert Sugden. A theory of focal points. *The Economic Journal*, pages 533–550, 1995.

Robert Sugden. *The Community of Advantage: A Behavioural Economist's Defence of the Market*. Oxford University Press, 2018.

Appendix A. - Experimental Instructions

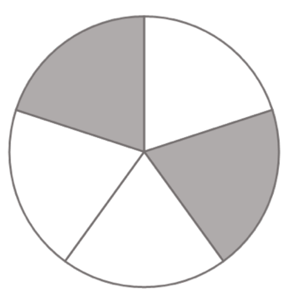
*Note: The instructions for the **I** treatment and the **G** are similar. When there are differences, the texts for the **I** treatment are put in curly brackets '{}', and the texts for the **G** treatment are put in square brackets '[]'.*

Experimental Instructions

Welcome to this experiment in [team] decision making. [During the experiment, you will be part of a team made of two people - you and another participant.] In each round of the experiment {you} [your team] will be matched with another {person} [team (also made of two people)] in the room. Your earnings will depend both on your [team's] decision and the decision of the other {person} [team]. There are 11 rounds (displayed in a random order) and the decisions you have to make depend on the type of task in that round. There are three types of task in the experiment.

TASK A

In this type of task, {you} [your team] and the other {person} [team] will be asked to choose one of the five slices of a pie, like the one shown below, by clicking on your choice. The pie is two-sided. The computer will flip and spin the pie so that [although] the side and rotation you see [will be the same as what your team-mate sees, they] may not be the same as the ones the other {person} [team] sees.



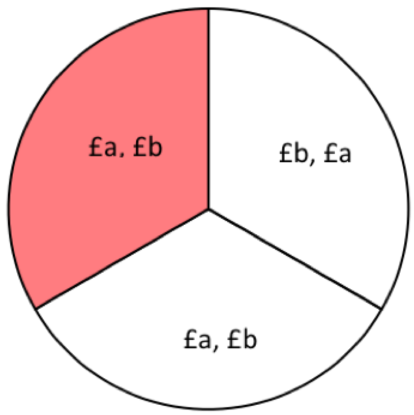
if you and the other {person} [team] choose the same slice each {of you} [player in the teams] will earn £5. If {you} [your team] and the other {person} [team] chooses

a different slice, {both of you} [you all] will earn nothing.

The experimenter has a pie made of paper. Raise your hand if you want to inspect it.

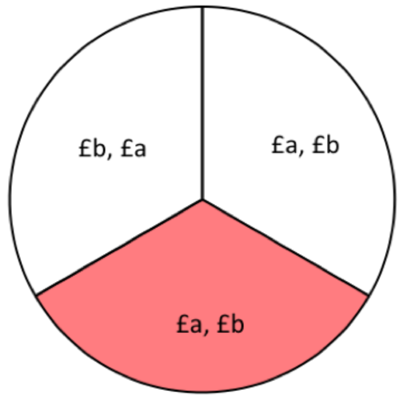
TASK B

In this type of task, {you} [your team] and the other {person} [team] will be presented with {a} [the same] pie with three slices, like the one shown below, and asked to choose one slice by clicking on your choice.



In each slice there are two amounts, represented by the letters 'a' and 'b' in the pie above. If {you} [your team] and the other {person} [team] **choose the same slice {you} each member of your team] will earn the amount on the left of the comma of the chosen slice** {and} [while each member of] the other {person} [team] will earn the amount on the right. If you and the other {person} [team] **choose a different slice you {both} [all] earn nothing in that task.** In the actual experiment the letters will be replaced by numbers.

The orientation of the pie is randomly decided by the computer separately for each {person} [team] in the room. This means that, although {you} [your team] and the other {person} [team] will see the same pie, its orientation will vary. For example, {you} [your team] may see the pie above while the other {person} [team] may see the pie below. There is therefore no way for you to know what orientation the pie of the other {person} [team] sees.



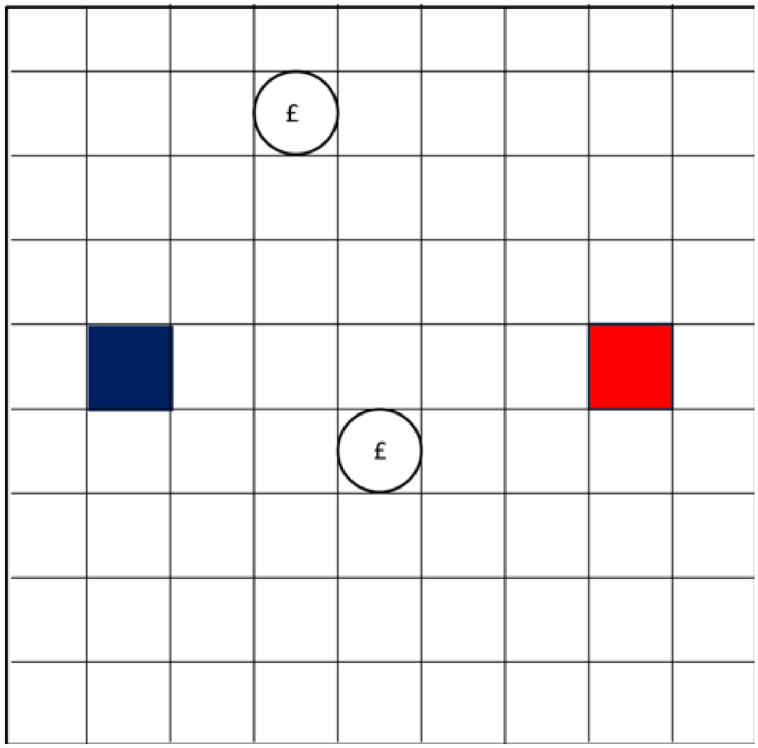
[NOTE. The amounts displayed are the earnings per player NOT per team!]

TASK C

In this type of tasks, you will be presented with a table similar to the one shown below. {You} [Your team] and the other {person} [team] will be assigned a colour. If {you are} [your team is] the Red {player} [team] then the other {person} [team] is the Blue {player} [team] and vice-versa. {You} [Your team] will be assigned a base, represented by a red square if {you are} [your team is] the Red {player} [team] or a blue square if {you are the} [your team is a] Blue {player} [team]. You will be told in each of these tasks whether {you are} [your team is] the Blue {player} [team] or the Red {player} [team].

There are two discs on the table. {You} [Your team] will be asked to choose one of these discs; you can do so by clicking on it. You earn the amount written inside the disc if the other {person} [team] has chosen the other disc. So, **you earn some money if {you and the other person} [both teams] choose a different disc. If {you} both [teams] choose the same disc you [all] earn nothing.**

[NOTE. The amounts displayed are the earnings per player NOT per team!]



[How to reach a team's decision

In every task, you will be asked to make a suggestion. Once you have made and submitted your suggestion you will be shown both what you have suggested and what your team-mate has suggested. These suggestions may either be the same or different. You will be asked to chat with your team-mate before you can proceed. A team's decision will be reached only if both you and your team-mate make the same suggestion and confirm it. You can revise your suggestions up to 5 times. If by then you and your team-mate have not reached the same suggestion, the computer will choose randomly for your team.]

Final earnings

At the end of the experiment the computer will randomly choose one of the 11 rounds. Your earnings will be determined (as explained in the previous pages) by {your choice} [the choice of your team], the choice of the other {person} [team] and the type of task played in the randomly chosen round. In addition to whatever you have earned in that round, you will be given a show up fee of £5.

Appendix B. - Level-K predictions

There are several models of level-K reasoning (Stahl and Wilson, 1995; Camerer et al., 2004; Costa-Gomes and Crawford, 2006). A common feature of these models is that players are assumed to maximise their own payoffs, and they are heterogeneous in terms of depth of reasoning. The lowest level, Level 0 ($L0$ henceforth) players do not play strategically. In some models, they are assumed to choose at random; in other models, they are assumed to choose the option that favours themselves. Higher-level players form beliefs about what players who are one level below them would do and best respond to that behaviour.

In this appendix, we demonstrate that in coordination games involving conflict of interest, group behaviour predicted by Level-K theory highly depends on the assumptions on the distribution of players' depth of reasoning (i.e. levels), the size of the conflict (i.e. difference in payoffs), and the assumption on $L0$ players' non-strategic behaviour. For simplicity, we first assume that $L0$ players will choose the option with the greater payoff for themselves with probability $p > 1/2$ (Crawford et al., 2008, used in). This assumption helps us demonstrate how prediction changes with p, the size of the conflict, and distribution of players' depth of reasoning. We will later discuss how different assumptions on $L0$ behaviour will affect the prediction.

Let us define Player 1 as the player favoured by the label cue (i.e. the player who has the higher material payoffs than their co-player in the focal point equilibrium) and Player 2 as the player unfavoured by it (i.e. the player who has the lower material payoffs than their partners in the focal point equilibrium). Based on this assumption, a $L0$ Player 1 will choose the red slice ($S1$) in the pie game and the close disc (C) in the bargaining table with probability $p > 1/2$, while a $L0$ Player 2 will choose $S2$ or the far disc (F) in the pie and bargaining game respectively with the same $p > 1/2$. A $L1$ Player 1's best response depends on both p and the difference in payoffs a and b (see 2). Other things being equal, if the difference is small or p is large, a $L1$ Player 1 will choose $S2$ and a $L1$ Player 2 will choose $S1$. $L2$ behaviour is the same as an $L0$, $L3$ behaviour is the same as an $L1$ and so and so forth. If the difference in payoffs is big or p is small, $L1$ players will switch their behaviour and choose $S1$ if they are Player 1 and $S2$ if they are Player 2. The same applies to higher levels. Therefore, Level-K predictions depend critically on the size of the conflict and p.

Notice that specifying these two parameters is still not enough to give clear predictions on groups' behaviour. When we move from individuals to groups, we assume that groups' depth of reasoning will be on the aggregate higher [8]. However, without specifying the distribution of levels, we are unable to know how the distribution of odd and even level players will change and this is critical to derive meaningful predictions.

Even if we were able to specify the distributions of levels in the population, Level-K predictions will still be sensitive to the assumptions on $L0$'s behaviour. Consider the Level-K theory proposed by Crawford et al. (2008) as an example. The model assumes that $L0$ only exists in the mind of the players and that the frequency of players with a level higher than 2 is negligible. As shown above, $L0$ players' non-strategic behaviour are specified as choosing the option with the greater payoff for themselves with probability $p > 1/2$. For simplicity, let's assume $p = 1$, which means $L0$ always choose the option that favours themselves. Under these assumptions, $L0$ Player 1 will choose the salient option (i.e. $S1$ or close disc)[9] , $L1$ Player 1 choose the non-salient option that favours their partners (i.e. $S2$ or far disc), and $L2$ Player 1 will choose the same as $L0$ Player 1. Accordingly, the model predicts that Player 1 in the **G** treatment are more likely to choose the label salient option than Player 1 in the **I** treatment. This is because the number of $L2$ players increase in the **G** treatment. However, if $L0$'s behaviour is specified as choosing a random option, as some Level-K models do, then $L0$ ($L2$) players in the new specification will behave the same as $L0$ ($L1$) in the old specification. Accordingly, the prediction of $L2$'s behaviour in the **G** treatment will be different, depending on which assumption is adopted on $L0$'s behaviour.

[8]Penczynski (2016) shows that a when low-level reasoning individual is paired with an individual with a higher level of reasoning, the former one tend to adopt a high level of reasoning because persuaded by the sophistication of their partner's argument.

[9]To explain subjects' behaviour in pure coordination pie games, level-K (Crawford et al., 2008) assumes that $L0$ players bias towards the red slice when payoffs are indifferent. We also keep this assumption so that $L0$ Player 1 will choose $S1$ rather than NS.