

# Testing team reasoning: Group identification is related to coordination in pure coordination games

James Matthew Thom\* Uzma Afzal† Natalie Gold‡

## Abstract

Games of pure mutual interest require players to coordinate their choices without being able to communicate. One way to achieve this is through team-reasoning, asking ‘what should we choose’, rather than just assessing one’s own options from an individual perspective. It has been suggested that team-reasoning is more likely when individuals are encouraged to think of those they are attempting to coordinate with as members of an in-group. In two studies, we examined the effects of group identity, measured by the ‘Inclusion of Other in Self’ (IOS) scale, on performance in nondescript coordination games, where there are several equilibria but no descriptions that a player can use to distinguish any one strategy from the others apart from the payoff from coordinating on it. In an online experiment, our manipulation of group identity did not have the expected effect, but we found a correlation of .18 between IOS and team-reasoning-consistent choosing. Similarly, in self-reported strategies, those who reported trying to pick an option that stood out (making it easier to coordinate on) also reported higher IOS scores than did those who said they tended to choose the option with the largest potential payoff. In a follow-up study in the lab, participants played either with friends or with strangers. Experiment 2 replicated the relationship between IOS and team-reasoning in strangers but not in friends. Instead, friends’ behavior was related to their expectations of what their partners would do. A hierarchical cluster analysis showed that 46.4% of strangers played a team reasoning strategy, compared to 20.6% of friends. We suggest that the strangers who group identify may have been team reasoning but friends may have tried to use their superior knowledge of their partners to try to predict their strategy.

Keywords: coordination, group identity, shared identity, social distance, team reasoning

---

\*Kantar Public, 4 Millbank, London, SW1P 3JA, UK. Email: james.thom@ipsos.com.

†Center for Behavioral Institutional Design, New York University Abu Dhabi, Saadiyat Marina District, Abu Dhabi, United Arab Emirates. Email: uzma.afzal@cbid.org.

‡Corresponding author. Centre for Philosophy of Natural and Social Sciences, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK. Email: natalie.gold@rocketmail.com.

# 1 Introduction

Game theory seeks to explain social interactions between individuals in terms of their interests. There is a vast literature on how people *cooperate* when their interests are partly coincident and partly opposed (e.g., Balliet et al., 2009; 2014; Sally, 1995). In these dilemmas, players should settle on the *Nash equilibrium*, a set of strategies where no player can benefit by changing strategy while the other players keep theirs unchanged. Sometimes, however, players' preferences for outcomes are perfectly aligned, and they must *coordinate* their choices without being able to communicate. Consider two people who are separated from each other in the London crowds and who want to reunite; each must decide independently to go to the same place. If all they care about is being reunited, then their interests are aligned and there will be multiple Nash equilibria — any pair of strategies leading to an outcome where they both go to the same place is a Nash equilibrium, no matter what the place is. Such situations are represented in game theory by games of *pure mutual interest* (Schelling, 1960).

Coordination poses a problem for classical game theory, which offers no explanation for how players choose between multiple Nash equilibria, even when one is associated with a higher payoff for both players. Consider a 2x2 Hi-Lo game (Figure 1). This is a pure coordination game with two Nash equilibria: (Hi, Hi), with payoffs of 2 for each player, and (Lo, Lo), with payoffs of 1. Both (Hi, Hi) and (Lo, Lo) are Nash equilibria because any one player changing strategy will result in a reduced payoff (in this case 0). Despite this, it seems intuitively obvious that people should play Hi, and the little experimental evidence we have confirms that virtually all players do so (Bardsley et al., 2010). The problem is that classical game theory vests all agency in individuals. Standard individualistic reasoning has players asking 'What should I do?', the answer to which is conditional on what the other player will do: if Player 1 chooses Hi then Player 2 should choose Hi, but if Player 1 chooses Lo then Player 2 should also choose Lo. Classical game theory therefore has no way of uniquely recommending or predicting the play of Hi.

**Team reasoning** explains coordination by allowing groups to act as agents, if individual players identify with the group (Bacharach, 2006; Colman & Gold, 2017, Gold & Colman, 2018; Sugden, 1993). Players engaging in team reasoning can ask 'What should *we* do?'

---

This work was supported by funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007–2013) / ERC Grant Agreement n. 283849. The experiments were conducted at Royal Holloway University of London EconLab and we are grateful for the support provided by Bjoern Hartig. Uzma Afzal gratefully recognizes financial support by Tamkeen under the NYU Abu Dhabi Research Institute Award CG005.

NG and JT conceptualised the experiments, JT implemented the experiments, analyzed the data, and drafted the experimental results. UA performed the hierarchical cluster analysis and drafted the section that reports it. NG drafted the remaining sections. All authors revised the drafts, approved the version to be published, and agreed to be accountable for all aspects of the work.

Copyright: © 2022. The authors license this article under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/).

FIGURE 1: Coordination game payoff matrix for a 2x2 Hi-Lo game.

	Hi	Lo
Hi	2, 2	0, 0
Lo	0, 0	1, 1

In the 2x2 Hi-Lo game above, the answer is clearly to choose the equilibrium associated with the highest payoff, also called the *payoff-dominant* equilibrium.

Team reasoning can explain coordination on payoff dominant equilibria, but so can cognitive hierarchy theory, also known as level- $n$  reasoning (Stahl & Wilson, 1994; Bacharach & Stahl, 2000; Camerer et al., 2004). This postulates a hierarchy of reasoners at different levels; each player assumes that their own strategy is the most sophisticated and plays a best response to a player who is one level lower in sophistication. If one fixes the strategy of the lowest level-0 reasoners, then it is possible to derive the strategies of higher-level players. In early applications, level-0 players were modeled as randomising over strategies with a uniform distribution (Stahl & Wilson, 1994; Bacharach & Stahl, 2000; Camerer et al., 2004). However, nothing in the basic theory requires that, so, for instance, a later application models level-0 players as trading off ‘payoff salience’ vs. ‘label salience’ (Crawford et al., 2008). For now, when we refer to cognitive hierarchy theory we will mean the early version where level-0 players randomise, but we will return to more sophisticated versions in the discussion.

To compare team reasoning and cognitive hierarchy theory, experimenters have used ‘nondescript’ Hi-Lo games, which contain several equilibria with differing payoffs, but there are no descriptions that a player can use to distinguish any one strategy from the others apart from the payoff from coordinating on it. Schelling (1960, pp. 295–296) describes a coordination game with four available strategies: (10, 10, 10, 9). Each payoff is associated with a distinct label, but players cannot see this label and therefore cannot use it to coordinate. Imagine, for instance, that the players are choosing between cards with a number on one side (the payoff if they manage to coordinate on the card) and a pattern on the other (the label). The cards are number side up, so the players cannot see the pattern, and the cards are shuffled separately for each player, so they cannot coordinate using the position of the cards. Thus, while 10 is the better payoff, there are three cards that say 10 with no obvious way of telling them apart, and the players only get the payoff if they choose cards that have the same label on the underside.

While cognitive hierarchy theory will always recommend choosing the highest paying option (if level-0 players randomize), team reasoning can explain why the solution may be for the pair to choose (9, 9). According to Bacharach and Bernasconi (1997), a player can choose between ‘options’ as she describes them only in a ‘framed game’ (see Sugden, 1995, for a complementary way of operationalizing this idea). So, in this game, she can either pick at random, pick a ten or choose the nine. Two players who both pick a ten stand a 1/3 chance of coordinating, with an expected payoff of 3.3. But the card with a nine on the front

is unique. If both players choose that card then they will get an expected payoff of nine. A team reasoning player chooses the option that is best for the group. Both players getting an expected payoff of 9 is better than an expected payoff of 3.3, so they should choose the 9-card.

Experiments comparing cognitive hierarchy theory and team reasoning in nondescript coordination games have had mixed results; even within individual experiments, choices in some games can only be explained by team reasoning theory and some can only be explained by cognitive hierarchy theory (Bardsley et al., 2010; Colman et al., 2014; Crawford et al., 2008; Faillo et al., 2016; Isoni et al., 2013; Mehta et al., 1994). Therefore, we sought an experimental design that could provide direct evidence about team reasoning in coordination games.

Team reasoning, but not cognitive hierarchy theory, predicts that group identification will be associated with coordination. Although these two main theories of team reasoning differ in the details of what group identification involves, the two theories both imply that increased group identification should lead to increased team reasoning (Gold, 2012). If we can promote group identification by ‘we-priming’, then people are more likely to think of a situation as posing a problem for ‘us’ and therefore more likely to use team reasoning.

The basic idea of team reasoning, as spelled out above, is that a team reasoner asks ‘what should we do?’. However, answering it by maximising expected payoffs is relatively *sophisticated*. Faillo et al. (2016) contrast that with *naïve* team reasoning, where players rely on some other heuristic to coordinate. For example, a naïve team reasoner might be motivated to simply ‘pick the odd one out’. They have framed the choice as a team reasoner should – “what should *we* choose?” – but do not opt for the payoff with the highest expected value.

In many cases, a sophisticated team reasoner will choose the same payoff as a naïve team reasoner. Consider the set: {10, 10, 10, 10, 10, 9}. In this set, any team reasoner should choose 9. Sophisticated team reasoners will identify 9 as the maximum expected value available to the pair. Naïve team reasoners, however, could choose 9 because it stands out as the unique payoff. Other sets can discriminate between the two types of team reasoner. For instance, in the set {10, 10, 10, 10, 10, 1}, sophisticated team reasons will pick a 10 because it maximises expected value but a naïve team reasoner would choose the 1 because it is unique.

In this paper, we report two experiments that directly test for the presence of team reasoning in coordination games. The payoff sets in the second were explicitly designed to discriminate between sophisticated and naïve team reasoning. In Experiment 1 we attempted to manipulate group identity in an online experiment, using we-priming techniques with pairs of strangers. Experiment 2 was conducted in the laboratory, where we generated a difference in group identification between conditions by asking participants to bring a friend to the laboratory with them and then varying whether participants played with their friend or with a stranger.

A measure of group identification that is increasingly being used by experimenters is the Inclusion of Other in Self (IOS) scale (Aron et al., 1992) shown in Figure 2, which measures how close an individual feels to a specific other person.

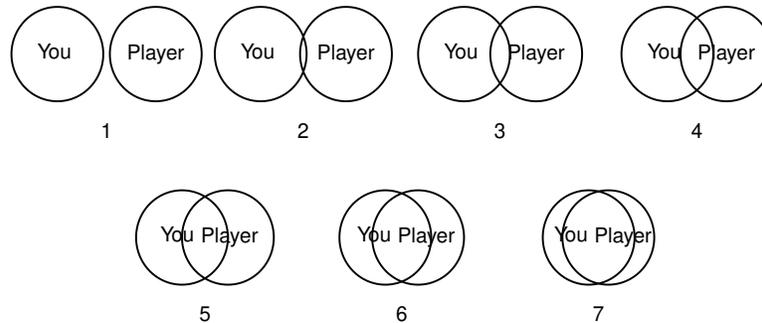


FIGURE 2: The Inclusion of Other In Self Scale. Respondents are asked to select the pair of circles that best describes the relationship between them and the other player.

IOS, considered as a psychological construct, is a consequence of group identification. Brewer and Gardner (1996, p. 84) argue that group identity is derived from a larger collective or social category and when we group identify ‘the boundaries of the self are redrawn, and the content of the self-concept is focused on those characteristics that make one a “good” representative of the group or of the relationship’. They explain that this involves a blurring of the boundaries between the self and a partner, in a manner that is captured by the IOS.

Then the IOS scale measures this blurring of boundaries. Aron et al. introduced the IOS in the laboratory in the context of romantic relationships, but Brewer and Gardner applied it to any social group from which we might derive a group membership. Gächter et al. (2015) validated the measure for less close relationships and for data gathered over the internet, showing that it is a psychologically meaningful and highly reliable measure of the subjective closeness of relationships.

IOS is related to cooperative behaviour in mixed motive games, both social dilemmas (De Cremer & Stouten, 2003) and weak-link coordination games (Gächter et al., 2017). It mediates Social Value Orientation (Cornelissen et al., 2011), which is a predictor of cooperation in social dilemmas (Balliet et al., 2009). The IOS scale is consistent with many different underlying processes and theoretical orientations. It is simply a measure of an individual’s subjective perception of their degree of closeness to another person.

Our aim in Experiment 1 was to manipulate group identity and in Experiment 2 to create variation in group identification between treatments by manipulating whether participants played with friends or strangers. We used IOS as a measure of that group identification. However, there were no treatment effects and our manipulation of group identity in Experiment 1 did not appear to have had the expected effect (see Section 4 for discussion). Nevertheless, we found a consistent correlation between group identification and team reasoning amongst strangers, and the results of Experiment 2 suggested that friends and strangers might be using different strategies. Therefore, we ran a post hoc hierarchical

cluster analysis, to investigate inductively the strategies that our participants were playing and to test the hypothesis that participants played different strategies with friends than with strangers.

## 2 Experiment 1

Experiment 1 consisted of three components: (1) a manipulation of group identity, (2) an incentivized two-person coordination game, and (3) a short questionnaire assessing understanding of the game, perceptions of group identity, and self-reported strategies. We had two overarching predictions in this experiment: (1) the manipulation of group identity should affect the tendency to coordinate using team reasoning, (2) there should be an overall relationship between our measure of group identity and the tendency to attempt to coordinate using team reasoning.

### 2.1 Methods

#### 2.1.1 Participants

Experiment 1 was completed by 632 participants. A further 56 participants were excluded from our analyses because they did not match the advertised recruitment criteria. Participants were recruited using Prolific Academic, an online recruitment panel. Eligible participants were all either British or American, were aged 18–30 years (mean age: 22.02 years), and were students. All participants received a £1.50 fixed participation fee, in addition to their winnings from the task (mean: £1.20; range: £0.00–7.95, including income from common-fate lottery – see below for details).

#### 2.1.2 Manipulation of group identity

There were two between-subject conditions, the *we* condition and the *they* condition, and in the coordination game participants were always matched with someone from the same condition as themselves. The manipulation of group identity preceded the coordination game and was comprised of several parts: A ‘Pronouns Task’ completed before the coordination game, matching pairs of participants on aspects of shared identity, a common-fate vs. individual lottery, and changes to the instructions of the coordination game.

The Pronouns Task required participants to count the number of pronouns in a short passage of text ([Supplementary Materials S1](#)). In the *we* condition, the pronouns were all in-group plural pronouns (e.g., *we*, *our*, *ourselves*, etc.), while in the *they* condition, the pronouns all referred to other people (e.g., *they*, *their*, *their selves*, etc.). The passages were otherwise identical, and in both cases contained 20 pronouns. This task (including the passage used) has previously been used to induce group identity on questionnaire measures (Brewer and Gardner, 1996).

Winnings from the coordination game were subject to a lottery, where they could potentially be trebled in value, which was either ‘common fate’ (*we*) or individual (*they*). After the experiment ended, the experimenter rolled a die to determine whether winnings would be trebled; in the *we* condition there was one die roll for the pair, in the *they* condition each participant was rolled for separately. If the experimenter rolled a 6, that pair/participant would have their winnings from the game trebled. Since partners in the *we* condition were rolled for as a pair, they both either won or lost together, unlike those in the *they* condition. Participants were already told about the structure of the lottery they would face in the instructions at the beginning of the experiment, so they already knew about it when they were playing the coordination games. The common fate lottery has previously been used to induce group identity in public goods games (Kramer & Brewer, 1984).

Participants in the *we* condition were matched so that they shared the same nationality and shared a political preference regarding the 2016 US Presidential race (Hillary Clinton/Bernie Sanders/Donald Trump/don’t care).<sup>1</sup> Participants in the *we* condition were reminded of these similarities, as well as the fact that they and their partner were both students, and both within the same age bracket. Participants in the *they* condition were simply told they would be matched with another participant in the experiment. Making natural shared identities salient has also previously been used to induce group identity in experiments (Castro, 2008; Dorrough et al., 2015; Fiedler et al., 2018; Jackson, 2008; Kiyonari & Yamagishi, 2004; Perdue et al., 1990; Platow et al., 1999; Yamagishi et al., 2005).

Finally, the instructions for the coordination game differed by condition in how they referred to the person with whom the participant had been matched. In the *we* condition, this was ‘your partner’, while in the *they* condition, it was ‘the other person’ (Full instructions can be found in [Supplementary Materials S2](#)). Again, using *we*-language in the instructions has previously been used in public goods games (Cookson, 2000).

### 2.1.3 Coordination game

The coordination game consisted of a series of choices between options worth varying numbers of points (Table 1), which corresponded to real monetary gains (5p per point). If both participants chose the same option, they would both receive those points. If they chose different options, both would receive nothing on that question. Participants did not know the identity of their partners and were not provided with any opportunity to communicate or any feedback between choices. The payoff sets available in our coordination game were a modified version of those reported by Bardsley et al. (2010). Our team reasoning outcomes are ex ante Pareto-dominant, but not ex post, so it is a test of ‘sophisticated’ team reasoners, in the sense of Faillo et al (2016).

---

<sup>1</sup>These figures were chosen since they were likely to be well known to British as well as American participants. This experiment was carried out after Hillary Clinton and Donald Trump were selected as the Democratic and Republican candidates respectively, so our choice was for ‘in an ideal world’, and offered Bernie Sanders as an option.

The coordination game was presented as a series of choices between cards with points values printed on their front (e.g., Figure 3). In order to convey to participants that the options were different but nondescript, participants were told that each card on the screen had a unique back pattern, visible only to the computer, and that their job was to pick not just a card worth the same number of points, but to pick the exact same card. They were also told that the location of the cards from left to right on the screen was randomised independently for each partner.

TABLE 1: Payoffs available in each question of the coordination game. Each number corresponds to a different option, so where there are three ‘10’s, there were three options available that were worth 10 points each. For each question, the options we considered to be consistent with team reasoning are shaded grey.

Choice	Payoffs
1	10 10 10 10 10 9
2	10 10 10 9 9 6
3	10 10 10 10 9 9
4	10 10 10 10 10 10
5	10 10 10 9 9 9
6	10 10 10 10 10 1
7	10 10 10 9 8 7
8	10 10 10 9 9 8
9	10 10 10 9 8 8
10	10 10 10 9 9 3

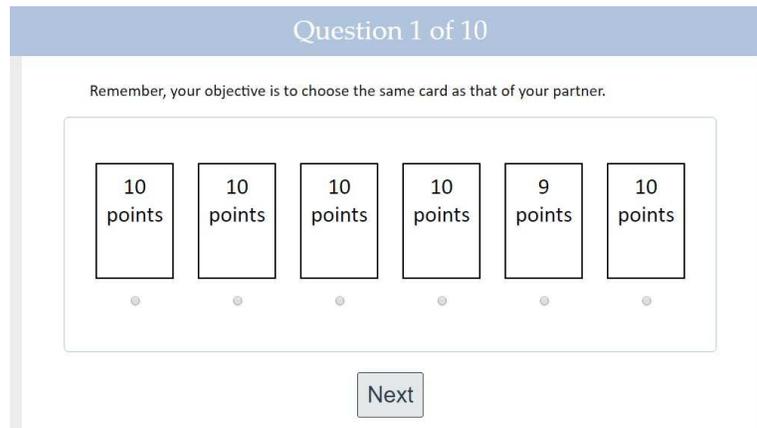


FIGURE 3: Screenshot of the coordination game, as seen by participants. In this case, the participant is in the *we* condition since the instructions above refer to ‘your partner’ rather than ‘the other person’.

### 2.1.4 Questionnaire

After finishing the coordination game, participants were asked to answer several sets of questions about the game (for screen shots of these, see [Supplementary Materials S3](#)).

The first set assessed the strategies participants employed in the game. Participants were asked to choose, from a list of strategies, which one best described how they completed the task: *Random*, *Pick High* (always chose an option worth the most points), *Pick Low* (always choose an option worth the least points), *Location label* (chose based on position on the screen), and *Payoff Label* (chose the option that stood out in terms of points value). (Participants saw the descriptors, not the strategy names.) Of the five strategies, we considered *Payoff Label* to be most suggestive of team reasoning (because team reasoning was developed to explain the rationality of players doing their part in payoff dominant outcomes), although the strategy is not a good descriptor for all the team reasoning choices in our experiment. Participants then completed the ‘Win-Points Scale’, a 5-point scale that asked participants whether they thought more about choosing the same card as their partner ‘winning’, or maximizing the number of points they would win if they did happen to choose the same card, ‘points’ (1 ‘only winning’ to 5 ‘only points’).

The second set asked how well participants understood the game, (1) on a 5-point scale, and (2) whether they understood that they needed to choose the same card as the other player, with three possible answers: ‘No’, ‘Yes, but I had no idea how to do that’, ‘Yes, and I think I had an idea of how to do that’.

The third and final set assessed participants’ beliefs about their partners. Participants were asked how similar they thought their partners would be to their selves, as a check for the effectiveness of our manipulation. They were also given the ‘Inclusion of Other in Self’ (IOS) scale, which has been used as a measure of shared identity or ‘oneness’ (Aron et al., 1992; Gächter et al., 2015). It is this second, deeper measure of shared identity that we expected to relate most closely to coordination behaviour. Finally, participants were asked how well they thought their partners understood the task.

### 2.1.5 Scoring

Each response to every question was recorded for analysis. Additionally, a composite measure was used to quantify participants’ choices across the questions. *Team Reasoning Score* was the number of questions on which the participant made a choice that was compatible with a team reasoning approach.

### 2.1.6 Hypotheses

- H1. IOS scores to be higher in the *we* condition than in the *they* condition.
- H2. Team Reasoning Scores to be higher in the *we* condition than in the *they* condition.
- H3. Participants in the *we* condition to be more likely to report using a *Payoff Label* strategy (which is consistent with team reasoning), than those in the *they* condition (who we expected

to favor *Pick High*).

H4. Positive overall correlation between IOS scores and Team Reasoning Scores.

H5. Higher IOS scores among participants who reported using a *Payoff Label* strategy, than among those who reported using a *Pick High* strategy.

## 2.2 Results

### 2.2.1 Understanding of the task

Three participants reported that they did not understand the task and were excluded from all subsequent analyses. Of the remaining participants, self-reported understanding of the task on the Likert scale did not differ by condition, *we*: 4.18(SE:  $\pm 0.044$ ), *they*: 4.24( $\pm 0.043$ );  $t(627) = -.95$ ,  $p = .344$ .

### 2.2.2 Pronouns task

In the Pronouns Task, participants reported an average of 18.43( $\pm 3.39$ ) pronouns (the correct answer was 20). Performance on the Pronoun task did not differ by condition, *we*: 18.35( $\pm 3.03$ ), *they*: 18.52( $\pm 3.74$ );  $t(583) = -.63$ ,  $p = .527$ .

### 2.2.3 Group identity manipulation

Contra Hypothesis 1, our manipulation did not have a detectable effect on group identity. Participants in the *we* condition did not report statistically significantly higher IOS than participants in the *they* condition (*we* = 3.28( $\pm 0.074$ ); *they* = 3.12( $\pm 0.083$ );  $t(614) = 1.43$ ,  $p = .154$ ; 95% CI of the difference:  $-.06$  to  $.38$ ). The conditions did differ in participants' estimates of how similar their partner was to their selves (*we*: 3.34( $\pm 0.04$ ), *they*: 3.14( $\pm 0.05$ );  $t(616) = 3.10$ ,  $p = .002$ ; Cohen's  $d = 0.25$ ), but this is unsurprising because participants in the *we* condition were explicitly told of ways in which they were similar to their partners. Our manipulation therefore did not appear to have the expected effect on perceptions of group identity. This conclusion is supported by Bayesian analysis: the Bayes factor for the lack of effect of our manipulation on IOS is 8.09 indicating that our result is 8.09 times more likely given the null hypothesis that the manipulations do not affect group identity than given the alternative hypothesis that they do; a Bayes Factor greater than 3 is generally considered 'substantial' evidence in favour of the null hypothesis (Rouder et al., 2009; Wagenmakers et al., 2011; Dienes, 2016).

### 2.2.4 Coordination game

Consistent with the observed lack of effect of our manipulation of group identity, there was no difference in Team Reasoning Score between the *we* and the *they* conditions (*we*: 4.82 $\pm 0.119$ ; *they*: 4.94 $\pm 0.13$ ;  $t(622) = -.69$ ,  $p = .492$ ; 95% CI of the difference:  $-.46$  to  $.22$ ). Neither did we find the predicted difference between conditions at the level of

individual questions. There were seven questions designed to differentiate team reasoning-based strategies from choosing the card associated with the maximum number of points (Questions 1, 2, 3, 7, 8, 9, and 10). In each question we predicted a higher proportion of team reasoning-compatible choices in the *we* condition than in the *they* condition. In six of these seven questions, the opposite was true (for a full breakdown, see [Supplementary Materials S4](#)). In Question 10, the difference in distribution of responses between the two conditions was not statistically significant ( $\chi^2(1) = 0.95, p = .331$ ). Overall, contra Hypothesis 2, there is no evidence of an effect of condition on participants' choices in the coordination game.

### 2.2.5 Reported strategy and coordination

The most common reported strategies were *Payoff Label* (47.4%.) and *Pick High* (32.3%). Contra Hypothesis 3, respondents in the *we* condition were not more likely to report using a *Payoff Label* strategy than those in the *they* condition.<sup>2</sup>

### 2.2.6 Inclusion of Other in Self and coordination

Because there were no clear effects of condition on choices in the coordination game or on reported strategy, participants from the two conditions were pooled for subsequent analysis of the relationship between group identity and coordination.

There was a positive overall correlation between IOS score and Team Reasoning Score as predicted by Hypothesis 4 ( $r(629) = .18, p < .001$ ). This result was robust to exclusion of participants who did not report using the two most common strategies, *Pick High* and *Payoff Label* ( $r(501) = .16, p = .001$ ), as well as exclusion of participants who did not have a clear understanding of how to successfully coordinate in the task ( $r(404) = .13, p = .012$ ). Given the correlation between IOS and Partner-Similarity, it is unsurprising that Partner-Similarity is also correlated positively with Team Reasoning Score ( $r(629) = .15, p < .001$ ). Partial correlation analysis finds the positive relationship between IOS and Team Reasoning Score is reduced only a little even when Partner-Similarity is included ( $r(626) = .13$ , as opposed to .18).

The relationship between IOS and Team Reasoning Score also held at the level of individual questions. Questions 1, 2, 3, 7, 8, 9 and 10 all offered options that could distinguish between a *Pick High* strategy and team reasoning. In each of these questions, mean IOS of those participants who chose a team reasoning-compatible option was numerically higher than those who did not, and these differences were statistically significant in six of seven of the questions. Taken together, these results are suggestive of some relationship between group identity and team reasoning.

<sup>2</sup>*We*: 46.2%, *they*: 48.7%. A Chi-Square test found that the numeric difference – which was in the opposite direction to that expected – was not statistically significant ( $\chi^2(1) = 0.40, p = .525$ ). See [Supplementary Materials S5](#) for frequencies of these two reported strategies and a breakdown against choices in the coordination game.

We expected that participants' reported strategies, as well as their actual choices, would be related to perceived group identity, in Hypothesis 5. Consistent with this expectation, IOS scores were higher on average for self-reported *Payoff Label* participants,  $3.40(\pm 0.08)$ , than for *Pick High* participants,  $3.15(\pm 0.95)$ , and this mean difference was statistically significant ( $t(441) = 2.02, p = .044$ ; Cohen's  $d = 0.19$ ). In addition to the mean differences in IOS and Team Reasoning Scores between strategies, there was also a positive correlation between the two factors in *Payoff Label* participants ( $r(298) = .18, p = .001$ ).

### 2.2.7 Summary

Experiment 1 found evidence of a positive correlation between IOS and Team Reasoning Score ( $r(629) = .18, p < .001$ ). Both Team Reasoning Scores and IOS were highest among participants reporting a *Payoff Label* strategy, and the correlation between the two factors was preserved in this group ( $r(298) = .18, p = .001$ ). Taken together, these results support our hypothesis of a relationship between shared identity and team reasoning (H4 and H5). However, because our treatment did not appear to affect perceived shared identity, we cannot make causal inferences (and there was no support for H1, H2, or H3). In order to provide a replication of our results and to test causality, we ran a second experiment with a different manipulation designed to manipulate IOS between treatments.

## 2.3 Discussion of null treatment effects

Our manipulation in Experiment 1 had no effect on group identity, as measured by the inclusion of self in other (IOS) score. Indeed, Bayesian analysis shows that our lack of effect actually provides evidence for the null, that the manipulations did not affect IOS. This was surprising, given the longstanding and influential theory of 'minimal groups', which holds that even assigning subjects to groups based on a trivial task will affect behaviour (Tajfel and Turner, 1979; 1986). However, the null effect of some of our treatments may be less surprising in the light of the current 'replication crisis' across experimental sciences and social sciences. Many of our manipulations were basically priming tasks and there is accumulating evidence that many priming results are false positives (Cesario, 2014; Harris et al., 2013; Pashler et al., 2013; Rohrer et al., 2015; Vadillo et al., 2016). In fact, the general evidence against the efficacy of priming leads us to be skeptical about the robustness of the minimal group paradigm. Other experimentalists have also found no effect of minimal group manipulations, although stronger group manipulations have sometimes been effective (Chen & Chen, 2011; Eckel & Grossman, 2005; Guala et al., 2013).

However, some of the manipulations we used were not just priming tasks. We put subjects into pairs based on their nationality and political preferences. Other studies have shown that matching subjects according to nationality can induce national in-group favouritism (Fershtman et al., 2005; Fiedler et al., 2018); further, assignment by nationality can change beliefs concerning the other player as well as social preferences towards them

(Dorrough & Glöckner, 2016; Fiedler et al., 2018). It may be that our group assignment changed players' beliefs about their co-player in a way that worked against an effect.

Another potential problem is that we measured IOS at the end of the study. Others have found a strong effect of group identification manipulations when IOS was measured directly after the manipulation (Dorrough et al., 2015). It may be that, in the absence of any feedback during the play of the coordination games, the effect of our group identity manipulations diminished over the course of time. So there would have been no effect of group identity on behaviour because group identity decayed, not because the treatment had no effect.

An alternative explanation of the lack of an observed treatment effect on perceptions of shared identity concedes that minimal group manipulations are effective, if small (Balliet et al., 2014), but notes that these results have usually been found in mixed motive games. Maybe the mechanism by which minimal group manipulations work is one that is effective in mixed motive but not coordination games. Our priming manipulation did increase subjects' perception of their similarity to their partner. We are more positively disposed towards people who we perceive as being similar to us (Byrne, 1961). Hence, we might be more inclined to be generous or altruistic towards them. Therefore, perhaps a priming manipulation that increased perceived similarity could affect play in mixed motive games like public goods games, where altruism or generosity might affect play, but not in coordination games.

## 3 Experiment 2

Experiment 2 consisted of a two-person pure coordination game completed in the lab. The task was structured similarly to that presented in Experiment 1, with the exception that we varied group identity between conditions by pairing participants either with a friend or with a stranger.

We had three overarching predictions for Experiment 2: (1) participants paired with a friend should feel a greater sense of shared identity with their partners than those paired with a stranger, (2) this should result in more team reasoning among those playing with a friend, and (3) we expect to replicate the overall relationship between group identity and team reasoning that we observed in Experiment 1.

### 3.1 Methods

#### 3.1.1 Participants

Experiment 2 was completed by 220 participants (139 female – 63.2%), aged 18–32 years old (mean age = 19.6 years). Participants were recruited using a panel managed by Royal Holloway University London's EconLab, and testing took place in an EconLab computer room. All participants were asked to attend the experiment with a friend, so half of all

subjects were recruited directly from the panel, and the other half were brought by a friend. All participants received a £4 fixed participation fee, in addition to their winnings from the task.

### 3.1.2 Manipulation of group identity

Upon arrival, participants were assigned to a condition, which determined whether they would play the coordination game with the *friend* they brought to the experiment, or with a *stranger* in the same session. Condition assignment was determined by the order in which participants arrived at the lab, counterbalanced across sessions. Before taking their seats, all participants were required to stand next to their assigned partner and confirm that they understood they would be playing the game with this person. This ensured that all participants knew who their partner was, regardless of condition. All that varied was *how well* they knew this partner.

### 3.1.3 Coordination game

The coordination game differed from that in Experiment 1 in three ways: (1) The payoff sets were altered and the number of sets offered was increased from 10 to 20 (see Table 2), (2) participants knew who their partner in the game was, and (3) each point was worth 10p, instead of 5p. The instructions were essentially the same Experiment 1 (they referred to ‘your partner’ but were otherwise the same as the ‘they’ condition, see [Supplementary Materials S6](#)).

### 3.1.4 Background Questions

Participants were asked the same background questions as in Experiment 1 (see [Supplementary Materials S3](#) for screenshots). In addition, participants in the *stranger* condition were asked whether, by chance, they were already familiar with the partner they had been assigned. Participants in both conditions were asked whether they thought their partner was more concerned about choosing the same card as their partner ‘winning’, or maximizing the number of points they would win if they did happen to choose the same card, ‘points’ (1 ‘only winning’ to 5 ‘only points’), which we will refer to as the ‘Partner-win-points’ scale.

### 3.1.5 Scoring

Each response to every question was recorded for analysis. Additionally, a Team Reasoning Score was calculated for each participant, as in Experiment 1. The Team Reasoning Score was the number of questions on which the participant made a choice that was compatible with team reasoning (see Table 2).

TABLE 2: Payoffs available in each question of the coordination game in Experiment 2. The predicted choices of a team reasoner are shaded grey. Payoff sets are clustered by type, designed to contrast different strategies (see Section 4 on naïve vs sophisticated team reasoning). Participants answered questions in the order shown in the left-most column. This order was determined pseudo-randomly, with no two questions of the same type adjacent.

Order	Payoffs					
1	10	10	10	10	10	9
10	12	12	12	12	12	11
18	10	10	10	10	10	1
13	12	12	12	12	12	1
3	10	10	10	9	9	7
9	12	12	12	11	11	10
11	10	10	10	9	9	1
16	12	12	12	11	11	1
2	10	10	10	9	8	8
6	12	12	12	11	10	10
8	10	10	10	10	9	9
14	12	12	12	12	11	11
19	10	10	10	10	1	1
4	12	12	12	12	1	1
12	10	10	10	9	8	7
17	12	12	12	11	10	9
7	10	10	10	5	4	3
20	12	12	12	5	4	3
5	12	12	12	3	2	1
15	10	10	10	3	2	1

### 3.1.6 Hypotheses

Our hypotheses were the same as for Experiment 1:

- H1. IOS scores to be higher in the *friend* condition than in the *stranger* condition.
- H2. Team Reasoning Scores to be higher in the *friend* condition than in the *stranger* condition.
- H3. Participants in the *friend* condition to be more likely to report using a *Payoff Label* strategy (which is consistent with team reasoning), than those in the *stranger* condition (who we expected to favor *Pick High*).
- H4. Positive overall correlation between IOS scores and Team Reasoning Scores.
- H5. Higher IOS scores among participants who reported using a *Payoff Label* strategy, than among those who reported using a *Pick High* strategy.

## 3.2 Results

### 3.2.1 Exclusions

Eight participants were excluded from our analyses because they were assigned to the *stranger* condition but happened already to have known the person they were partnered with, prior to attending the session.

### 3.2.2 Understanding of the task

All participants reported that they had understood what they were required to do, and self-reported understanding of the task on the Likert scale did not differ by condition (*friend*: 4.26(±.07), *stranger*: 4.15(±.07);  $t(209) = 1.10$ ,  $p = .272$ ; Cohen's  $d = 0.15$ ).

### 3.2.3 Group identity

Participants in the *friend* condition reported significantly higher mean IOS with their partners than did participants in the *stranger* condition (*friend*: 4.53(±.12); *stranger*: 2.20(±.13);  $t(209) = 13.48$ ,  $p < .001$ ; Cohen's  $d = 1.86$ ). The same was true of estimated similarity (*friend*: 3.55(±.09), *stranger*: 2.70(±.08);  $t(208) = 7.15$ ,  $p < .001$ ; Cohen's  $d = 0.94$ ); and ratings of partners' understanding of the game (*friend*: 4.34(±.06), *stranger*: 3.59(±.07);  $t(210) = 8.14$ ,  $p < .001$ ; Cohen's  $d = 1.15$ ). (For a frequency chart showing the distribution of IOS estimates in Experiment 2, split by condition, see [Supplementary Materials S7](#), Figure 1.) Overall, the manipulation of group identity in this experiment appears to have been successful and Hypothesis 1 was supported.

### 3.2.4 Team reasoning and coordination

Choosing consistently with team reasoning appears to have been an effective way to coordinate. The average Team Reasoning Score for each pair was positively correlated with the number of items<sup>3</sup> that the pair agreed upon ( $r(105) = .42$ ,  $p < .001$ ), as well as with that pair's payoff from the task ( $r(105) = .35$ ,  $p < .001$ ).

Despite the success of our manipulation of group identity, there was no significant difference in number of choices of the team-reasoning option between the conditions (*friend*: 10.04±.47, *stranger*: 9.63±.40;  $t(201) = .67$ ,  $p = .503$ ). Nor did condition appear to affect success in the game (Table 3). (For a frequency chart showing the distribution of Team Reasoning scores in Experiment 2, split by condition, see [Supplementary Materials S7](#), Figure 2.) Taken together, these results do not indicate a noticeable effect of condition on choices in the coordination game, contra Hypothesis 2.

A full break down of choices by experimental condition is provided in [Supplementary Materials S8](#).

<sup>3</sup>An item is a particular instance of a payoff, explained in this task by the invisible pattern on the back of the card. To select the same item, a pair cannot both simply choose '10', they must choose the *same* '10'.

TABLE 3: Mean numbers of items agreed on by both partners, and mean payoffs (GBP), both split by condition.

	Friend	Stranger
Mean items agreed upon	4.63 (0.55)	5.24 (0.50)
Mean payoff (£)	4.45 (0.50)	4.96 (0.45)

### 3.2.5 Inclusion of Other in Self and coordination

Despite the absence of the expected effect of treatment group on coordination behaviour, we did replicate our finding from Experiment 1 of a positive correlation between IOS and coordination ( $r(212) = .15, p = .027$ ). This result was robust to exclusion of participants who did not report using the two most common strategies (*Pick High* and *Payoff Label*;  $r(164) = .18, p = .023$ ). Hypothesis 4 was supported.

This replication is surprising, since our condition manipulation *did* affect IOS, but not coordination. One possibility is that the relationship between IOS and coordination holds primarily for strangers, since none of the participants in Experiment 1 knew more than a few details about their partners. Indeed, there is a positive correlation between IOS score and Team Reasoning Score in the *stranger* condition ( $r(110) = .25, p = .008$ ), but not in the *friend* condition ( $r(102) = .07, p = .468$ ). This pattern was robust to exclusion of participants who did not report using the two most common strategies (*stranger*:  $r(86) = .36, p = .001$ ; *friend*:  $r(78) = .01, p = .906$ ). To test whether these results indicate a genuine difference between conditions, we carried out a factorial ordinal probit GLM with Team Reasoning Score as the dependent variable and IOS and Condition as predictors.<sup>4</sup> To increase power, the sample for this model excluded those participants who had explicitly reported a strategy of choosing payoffs randomly (*Random* and *Location label*, which was effectively choosing randomly since the location of each item was randomized independently for each participant). The results of the model are presented as Model 1 in Table 4. There is marginal evidence of an interaction between Condition and IOS, which may indicate a more positive relationship between IOS and Team Reasoning Score in the *stranger* condition than in the *friend* condition.

Overall, these results suggest a positive relationship between IOS and the tendency to make team reasoning-compatible choices. However, this relationship is only observed in participants in the *stranger* condition. The results of the *stranger* condition are a replication of the results of Experiment 1, which was carried out entirely using strangers. The absence of the same pattern in the *friend* condition, however, is puzzling.

<sup>4</sup>A probit link function was used because the distribution of the dependent variable appeared non-normal. Our results do not differ from those obtained with a linear model.

TABLE 4: Ordinal Probit GLM output for *Payoff Label*, *Pick High*, and *Pick Low* participants. Coefficients shown for condition dummy variable are for the *friend* condition, using the *stranger* condition as a reference.

Dependent Variable: Team reasoning score	Model 1		Model 2	
	$\beta$ (SE)	p	$\beta$ (SE)	p*
Condition[friend]	0.483 (0.518)	.350	1.760 (0.765)	.021
IOS	0.229 (0.082)	.056	0.234 (0.082)	.123
Condition[friend]*IOS	-0.215 (0.128)	.092	-0.270 (0.129)	.037
Partner-win-points	–	–	-0.012 (0.123)	.025
Condition[friend]* Partner-win-points	–	–	-0.364 (0.173)	.035
-2 Log Likelihood	306.811		512.467	
Chi-Square (df)	8.00 (3)	.054	13.97 (5)	.016

\* Please note that it is not appropriate to interpret main effects in a model that has statistically significant interaction effects and for that reason the main effect is not discussed in the main text.

### 3.2.6 Reported strategy and coordination

Participants' self-reported strategies corresponded to their actual choices in the coordination game. As in Experiment 1, the most common reported strategies were *Payoff Label* (43.9%) and *Pick High* (33.5%). The *Payoff Label* participants had higher Team Reasoning Scores than the *Pick High* participants ( $PL: 13.62 \pm .38$ ,  $PH: 6.37 \pm .14$ ;  $t(162) = 16.01$ ,  $p < .001$ ; Cohen's  $d = 2.96$ ).

Contra Hypothesis 3, the distribution of all reported strategies did not differ by condition ( $\chi^2(4) = 2.28$ ,  $p = .685$ ), and neither did the distribution of the two most common strategies ( $\chi^2(1) = .005$ ,  $p = .942$ ). (For a frequency chart showing the distribution of reported strategy by condition in Experiment 2, see [Supplementary Materials S9](#).) Pairs in the *friend* condition were not more likely to report using same strategy than pairs in the *stranger* condition (39% and 50%, respectively).

Self-reports of strategies were consistent with participants' responses to the Win-Points scale. Participants identifying *Pick High* as their strategy scored higher on average ( $3.51 \pm .11$ ) than those identifying *Payoff Label* ( $2.54 \pm .09$ ;  $t(144) = 6.88$ ,  $p < .001$ ; Cohen's  $d = 1.14$ ).

We expected that IOS would be associated with the strategy adopted, as observed in Experiment 1. IOS scores were indeed higher on average for *Payoff Label* participants ( $3.47(\pm .18)$ ) than for *Pick High* participants ( $3.18(\pm .22)$ ), but this mean difference was not statistically significant ( $t(145) = 1.01$ ,  $p = .312$ ). As noted above, in Experiment 1 none

of the participants could know their partner, so we should primarily expect replications in the *stranger* condition. When the data are split by condition, *Payoff Label* IOS scores are indeed statistically significantly higher in the *stranger* condition (*Payoff Label*:  $2.39 \pm .22$ , *Pick High*:  $1.78 \pm .17$ ;  $t(84) = 2.08$ ,  $p = .041$ ; Cohen's  $d = 0.47$ ), but not in the *friend* condition (*Payoff Label*:  $4.68 \pm .16$ , *Pick High*:  $4.71 \pm .22$ ). Hypothesis 5 was only supported in the *stranger* condition and not in the experiment as a whole.

Reported strategies appear to correspond to the strategies that participants believed their partners had adopted. Of participants reporting a *Pick High* strategy, 88.7% believed their partners would use the same strategy as they did. The same was true of 65.6% of participants reporting a *Payoff Label* strategy.

Furthermore, we can compare participants' responses to the Win-Points scale to their placement of their partners on the same scale – the Partner-Win-Points scale. Win-Points and Partner-Win-Points were positively correlated ( $r(212) = .48$ ,  $p < .001$ ). This correlation was observed in both conditions (*friend*:  $r(102) = .63$ ,  $p < .001$ ; *stranger*:  $r(110) = .34$ ,  $p < .001$ ), but a Fisher r-to-Z transformation confirms that the Pearson coefficient is larger in the *friend* condition than in the *stranger* condition ( $Z = 2.83$ ,  $p = .005$ ).

The result above could offer a clue as to what is happening in the *friend* condition: perhaps these participants were primarily focused on predicting their partners' strategies, irrespective of their sense of shared identity. To test this, we expanded on Model 1 reported in Table 4 to include two additional predictors: Partner-Win-Points, and its interaction with the condition variable. There are two main findings from this model (Model 2): (1) there is further evidence for a condition\*IOS interaction, and (2) there is a statistically significant condition\*Partner-Win-Points interaction. Splitting the model by condition supports the idea that participants playing with friends chose according to different factors than those playing with strangers. In the *friend* condition, there was an effect of Partner-Win-Points but not IOS; in the *stranger* condition, there was an effect of IOS, but not Partner-Win-Points.

Model 2 (Table 4) appears to uncover a key difference between the conditions in how the coordination problem was approached. Participants playing with strangers had a generally low sense of group identity with their partners. In this group, an increase in IOS predicts an increase in Team Reasoning Score. By contrast, participants playing with friends had a generally higher sense of group identity. In this second group, an increase in IOS does not predict an increase in Team Reasoning Score. Instead, participants in the *friend* condition chose in accordance with their predictions of their partner's approach to the game.

It may seem surprising that Partner-Win-Points was not a statistically significant predictor in the *stranger* condition, given the requirements of the coordination game. This can perhaps be attributed to the effect that information about one's partner has on the two predictor variables. This idea is examined in detail in the Discussion. It should however be noted that beliefs about one's partner can be wrong. Indeed, there was no evidence of a positive correlation between a participant's placement of their self on the Win-Points scale, and where that participant's partner places them on the same scale, either overall ( $r(212) =$

-.10,  $p = .134$ ) or in the conditions separately (*friend*:  $r(102) = -.19$ ,  $p = .054$ ; *stranger*:  $r(110) = -.02$ ,  $p = .805$ ).

### 3.2.7 Summary

In Experiment 2 we replicated our main findings from Experiment 1: there is a positive correlation between IOS and Team Reasoning Scores in participants playing with strangers ( $r(110) = .25$ ,  $p = .008$ ; Hypothesis 4). Our IOS manipulation was successful (Hypothesis 1) and Experiment 2 also uncovered an apparent difference in approach between those partnered with a friend and those partnered with a stranger. Team Reasoning Scores were predicted by perceived shared identity in the *stranger* condition ( $r(110) = .25$ ,  $p = .008$ ) but not in the *friend* condition ( $r(102) = .07$ ,  $p = .468$ ). (Hypothesis 5 was supported only in the *stranger* condition.) In the *friend* condition, Team Reasoning Scores were only predicted by participants' expectations of what their partners would do ( $r(102) = .63$ ,  $p < .001$ ). Perhaps because of this, we did not observe the predicted overall difference in Team Reasoning Scores between the conditions (Hypothesis 2) or the predicted difference in self-reported strategies (Hypothesis 3). The implications of these findings for our understanding of coordination and team reasoning are discussed in the Discussion.

## 3.3 Cluster analysis of strategies: differences between friends and strangers

In this section, we explore the heterogeneity of strategies in our sample. Ex ante, we identified several strategies that participants might use in our games:

- Pick high (the strategy that would be advocated by cognitive hierarchy theory if level-0 players randomise; corresponds to our self-reported strategy of *Pick High*);
- Team reasoning (maximisation of ex ante expected value, could be considered sophisticated team reasoning; in terms of our self-reported strategies it might sometimes imply *Pick High* and other times *Payoff Label*);
- Pick the odd one out (potentially could be naïve team reasoning; in terms of our self-reported strategies, this is pure *Payoff Label*);
- Pick low (corresponds to *Pick Low*)
- Pick at random (corresponds to *Random*)
- Pick based on the position on the screen (corresponds to *Location*)

In Experiment 2, we specifically designed games that would distinguish these strategies, in particular Pick high, Team reasoning, Pick the odd one out, and Pick low. There were payoff sets where the unique payoff was not associated with maximal expected value available for a pair, e.g., {10, 10, 10, 10, 10, 1}; where the lowest payoff was neither unique nor associated with the maximum expected value, e.g., {10, 10, 10, 9, 8, 8}; sets where

the highest payoff, the payoff that maximised expected value and the lowest payoff all came apart, such as {10, 10, 10, 9, 9, 1} and {10, 10, 10, 9, 8, 7}. The full set of games is in Table 2.

One question we were interested in was whether participants who picked the unique payoff, even when it was not the one that maximised ex ante expected value, could nevertheless be considered to be doing naïve team reasoning. Three of our games (Exp 1, Q10; Exp 2, Q11, 16) had a payoff structure that could distinguish between the three most plausible ways that subjects could have made their decision: sophisticated team reasoning, picking one of the cards with the highest payoff, and choosing the card with the unique payoff. In each of these questions, those who chose the sophisticated team reasoning outcome reported higher mean IOS than those who chose either ‘pick high’ or ‘pick unique’ (Table 5). This difference suggests that picking the unique card, when that is not an ex ante maximising strategy, is associated with a reduced sense of shared identity relative to sophisticated team reasoning.

TABLE 5: Scores were compared using a one-way ANOVA. Mean IOS scores by choice strategy in three questions that distinguish between sophisticated team reasoning, picking high, and picking unique.

Experiment (question)	Mean IOS score by payoff choice strategy			Results of ANOVA
	Pick high (10/10/10)	Sophisticated team reasoning (9/9)	Pick low/unique (3)	
1 (10)	3.10 (1.32)	3.47 (1.33)	3.26 (1.48)	F(626,2)=2.44, p = 0.09
2 (11)	3.28 (1.77)	4.00 (1.41)	2.93 (1.63)	F(209,2)=3.94, p = 0.02
2 (16)	3.14 (1.76)	3.98 (1.41)	3.14 (1.76)	F(209,2)=3.95, p = 0.02

Following on from this, we derived a typology of strategies in our data using a hierarchical clustering technique (Fallucchi et al., 2019). Since games in Experiment 2 were explicitly designed to discriminate between strategies, we first identified strategies from that data and then used data from Experiment 1 as a model check, since we would expect the distribution of strategies in the Strangers condition in Experiment 2 to be similar to the distribution in Experiment 1.

### 3.4 Method

We created clusters using Ward's minimum variance method (Ward, 1963), which merges observations to create clusters that minimize the within-group variance (the sum of squares of the distances from the cluster mean). The number of clusters was determined using the Duda-Hart stopping rule, which is the sum of squared distances in the two groups to be clustered divided by the sum of squared distances within the cluster that would be created if they were combined, with larger values indicating more distinct clustering (Duda & Hart, 1973). For Experiment 2, given that participants displayed differences in approach across the two treatments, we examined decision strategies across the *friend* and *stranger* conditions separately.

### 3.5 Results

#### 3.5.1 Strategy clusters

In Experiment 2, in the *friend* condition, we observed three clusters, which seemed to correspond to the three or four main strategies we identified ex ante: Pick high, Team reasoning, and Pick low/ odd (see figure in [Supplementary Materials S10](#)). Given that picking low and picking odd often converged to the same choice for several questions, the third cluster seemed to represent both strategies combined. From a total sample of 102 participants in this treatment, 58.82% played the *Pick high* strategy, followed by *Team reasoning* and *Pick low/odd* strategies in equal proportions (20.59%). For participants in the *stranger* condition, we also observed 3 clusters displaying similar behaviour (see figure in [Supplementary Materials S11](#)). From a total sample of 110 participants, 46.36% had the strategy of *Team reasoning*, 40.0% played *Pick high*, and 13.64% played *Pick low/odd*. The difference in distribution of strategies between the two conditions was statistically significant,  $\chi^2(2) = 15.68, p = .000$ .

The clustering in Experiment 1 shows roughly the same distribution as the *stranger* condition in Experiment 2. From a total sample of 629 participants, 40.38% participants are in the *Team reasoning* cluster, 44.52% in *Pick high*, and 15.10% *Pick low/odd*. See figure in [Supplementary Materials 12](#).<sup>5</sup>

We can get an idea of how many in the third cluster were playing Pick Low rather than Pick odd by looking at choices in some of the individual games, which discriminate between the two strategies. There were two questions in Experiment 2 and one in Experiment 1 that discriminated between picking the lowest value and picking the unique value, with between 6.6% and 12.0% picking the lowest and between 39.2% and 48.6% picking the odd one out (see Table 6).

Finally, we also looked at mean differences in IOS scores across the three clusters for both Experiments, see Table 7. In both conditions for Experiment 2, as well as Experiment

<sup>5</sup>Note that for Experiment 1 the analysis was based on the pooled sample for both conditions.

TABLE 6: Breakdown of questions where the lowest payoff is neither unique nor associated with maximal expected value available for a pair. Each payoff is shown with the proportion of participants who chose underneath.

Experiment (question)	Percentage of subjects choosing item associated with each payoff		
1 (9)	Highest (10/10/10) 46.7%	Unique (9) 41.3%	Lowest (8/8) 12.0%
2 (2)	Highest (10/10/10) 51.9%	Unique (9) 39.2%	Lowest (8/8) 9.0%
2 (6)	Highest (12/12/12) 44.8%	Unique (11) 48.6%	Lowest (10/10) 6.6%

1, the *Team reasoning* cluster reported numerically higher mean IOS than those in *Pick high*. However, there were no statistically significant differences between the three means in either condition, as determined by one-way ANOVA for both Experiments (see Table 7).

TABLE 7: Mean IOS scores according to each of the three clusters obtained from hierarchical clustering analysis. The F-statistic shown is for the one-way ANOVA comparing the mean IOS scores across the clusters. Standard errors reported in brackets.

Condition	Mean IOS score for each cluster			Results of ANOVA
	Team reasoning	Pick High	Pick Low/Odd	
Exp. 2 Friend	4.57 (1.12)	4.52 (1.19)	4.52 (1.25)	F(2,99)=0.02, p=0.983
Stranger	2.25 (1.40)	2.11 (1.17)	2.31 (1.45)	F(2,115)=0.22, p=0.801
Exp. 1 None	3.32 (1.46)	3.08 (1.33)	3.23 (1.33)	F(2,626)=1.97, p=0.140

## 4 Discussion

We found evidence of a relationship between group identification and team reasoning behaviour in a coordination game. In Experiment 1, conducted completely anonymously over the internet, we found that there was a correlation between IOS score and team reasoning ( $r = 0.18$ ,  $p < .001$ ). In Experiment 2, there was less team reasoning behaviour amongst friends than amongst strangers, with a hierarchical cluster analysis finding that 46.4% of participants played a team reasoning strategy in the strangers condition, compared to 20.6% amongst friends. The distributions of strategies in the two conditions were statistically different and the results from the stranger condition was similar to the proportion of participants playing a team reasoning strategy in Experiment 1, 40.2%. We replicated the correlation between IOS and team reasoning in the stranger condition of Experiment 2 ( $r$

= 0.25,  $p = .008$ ). However, IOS did not correlate with team reasoning behaviour amongst friends ( $r = 0.07$ ,  $p = .468$ ). Instead, in the friend condition, team reasoning behaviour was correlated with subjects' expectations about their partners and the extent to which they thought their partners had thought more about choosing the same card as them versus maximizing the number of points they would win if they did happen to choose the same card ( $r = .63$ ,  $p < .001$ ). It seems that friends may have used their superior knowledge of their partner to try to predict their partner's strategy and then match it, while the relationship between IOS and behaviour among strangers is consistent with their using team reasoning, when they identified with their partner.

Previous experimenters also found a relationship between social cohesion and behaviour in conducted in mixed motive games (Attanasi et al., 2016; Gächter et al., 2017). They have found that group identification increases the likelihood of reaching the cooperative outcome, which would usually be associated with team reasoning. Indeed, Gächter et al. (2017) found that the effect of IOS in mixed-motive games was larger than the maximum effect that would be expected from a simple theory of other-regarding preferences. In contrast, in our pure coordination game, friends were *less* likely to play a team reasoning strategy. We did not find that friends did better than strangers at our task, nor was IOS associated with behaviour in the friends condition. Our results suggest that friends tried to predict their partner's behaviour and acted on that basis, but their predictions turned out to be no more successful than team reasoning. One possible explanation for the difference between findings in our coordination game and findings in mixed motive games is that IOS among friends may correspond to liking. In coordination games, liking and social preferences should not affect play (Karpus & Gold, 2016). However, they may affect the payoff orderings in mixed motive games, enabling friends to perform better in that environment.

Another possibility – unexplored here – is that a participant could take their friend's risk preferences into account when forming a strategy. The way we have operationalized sophisticated team reasoning implies that agents are maximizing expected value. Agents who are risk averse or who assume that others are risk averse may behave differently. This may help to explain the results in our friends condition, since those participants may have some knowledge of the risk preferences of their partner. To test this, future studies would need to explicitly assay participants' risk preferences as well as their beliefs about their partners' risk preferences. The implications of these considerations for team reasoning are unclear, as partners' preferences could inform either team reasoning strategies or individual best-response strategies. In either case, choosing would be expected to deviate from the sophisticated team reasoning that has been the focus of this paper.

Our finding that IOS is only predictive of behaviour among strangers (and not friends) is consistent with work comparing ingroup and outgroup behaviour in dictator games, where player's interests are perfectly opposed (Fiedler et al., 2018). Using nationality to create natural ingroups and outgroups, Fiedler et al. (2018) found that IOS was related to giving in dictator games that were played with an outgroup member, but not in games that

were played with an ingroup member, who shares cultural norms. As with our results, it is possible that IOS is predictive of behaviour in situations where people feel they lack relevant knowledge about their partners, but that when players have knowledge of their partners or of relevant cultural norms, then they use that instead, severing the relationship with IOS. This hypothesis could be tested in further research.

Our group identity treatment did not have the anticipated effect in Experiment 1 (see discussion in Section 2.3), which means that we produced no evidence that group identification causes IOS (though this is a theoretical connection that is often made in the literature), and our finding of a relationship between IOS and team reasoning in strangers was only a correlation. Nevertheless, the fact that friends had higher IOS than strangers and, yet, we did not find a relationship between IOS and team reasoning in the friends condition shows that it is not the case that team reasoning causes high IOS. It is still possible that team reasoning behaviour and high IOS are related in the stranger condition because they have a common cause, albeit a causal chain that is broken in the friends condition. However, it is hard to think of a convincing candidate for the common cause. For instance, IOS might measure a disposition to be cooperative, which could be a common cause of IOS and the choice of efficient strategies in mixed motive games. However, only coordination — and not cooperation — is involved in pure mutual interest games. Similarly, personality traits might be a candidate for a relationship between IOS and cooperation. Of psychologists' 'Big Five' personality traits (Goldberg, 1993), 'Agreeableness' is associated with cooperative behaviour in social dilemmas (Kagel & McGee, 2014; Volk et al., 2011; 2012), as is 'Honesty-Humility', a sixth personality trait that has recently been added to the model, which is associated with a prosocial Social Value Orientation and is a good predictor of prosocial behaviour (Hilbig et al., 2014).<sup>6</sup> Both Agreeableness and Honesty-Humility might plausibly be better predictors of behaviour in interactions with strangers — after all most people are agreeable and fair in their dealings with friends — and therefore explain why high IOS among strangers is associated with cooperation. However, it is hard to see why any of the big six traits would cause team reasoning (the other four are Openness to experience, Conscientiousness, Extraversion, and Neuroticism), especially in pure mutual interest games. Given the interesting connections, further research is needed on the relationship between personality traits, team reasoning, and IOS—in both mixed motive and mutual interest games.

We found three types of strategy that were consistently expressed: picking the highest number, picking the odd one out, and choosing the team reasoning solution. We wondered whether picking the odd one out was, in fact, a form of naïve team reasoning, where subjects are thinking “what should *we* do?”, but then using a simple heuristic rather than responding

---

<sup>6</sup>Honesty-Humility represents ‘the tendency to be fair and genuine in dealing with others, in the sense of cooperating with others even when one might exploit others without suffering retaliation’ (Ashton & Lee 2007, p. 156). So, while Agreeableness is associated with the tendency to cooperate even in the face of exploitation, people who are high in Honesty-Humility are less tempted to defect in order to exploit others in mixed motive games (Hilbig et al., 2013).

to the expected value of the items. However, picking the odd one out was not associated with a higher IOS, and it is plausible that those who consistently picked the odd one out were not using team reasoning at all, but just using the heuristic ‘pick a salient option’. Picking the odd one out has some similarities with the strategy that Mehta et al. (1994) call ‘primary salience’, which in turn is similar to a cognitive hierarchy theory where level-0 players only weight ‘label salience’ and not ‘payoff salience’ (Crawford et al., 2008). Label salience is salience due to attributes such as colour and position, which corresponds to picking the odd one out, whilst payoff salience tracks higher payoffs.

We did not find a difference in IOS score between the strategy clusters; results were numerically in the direction we predicted but were not statistically significant. However, we did find a relationship between IOS and participants’ self-reported strategies, with those who reported choosing the option that stood out in terms of points value (*Payoff Label*) having higher IOS than those who reported always choosing an option worth the most points (*Pick High*). The multiple choice we offered for the self-report questions doesn’t map perfectly onto the strategy clusters. Sophisticated team reasoning would sometimes involve picking the option that stood out, but sometimes it required picking the option with higher payoff, even when one with a lower payoff stood out. Therefore, it is hard to know which strategy team reasoners would have reported themselves as using. In addition, we took self-report measures at the end of the experiment, so they might reflect the strategy that participants used in the final games. In contrast, the clusters are derived from performance across all games. IOS was also measured after the games, potentially making it more likely to covary with the self-reports.

Potentially our data could also be explained by cognitive hierarchy theory, if we allow that different participants had different models of level-0 players. Some experimentalists have used a form of cognitive hierarchy theory where level-0 players trade off ‘payoff salience’ vs ‘label salience’ (Crawford et al., 2008). Payoff salience is assumed to be more strongly weighted than label salience. Therefore, as payoff differences increase, players who have the trade-off model of level-0 players will tend increasingly to choose higher payoff options. This gives three different ways of modelling level-0 players, which may correspond to the three strategies we found. Picking the highest number corresponds to a cognitive hierarchy theory where level-0 players randomise. Picking the odd one out corresponds to a theory where level-0 players only consider payoff salience. Potentially, trading off payoff salience and label salience could produce what we have thought of as the team reasoning strategy, though this remains to be shown via a formal model. However, cognitive hierarchy theory has no obvious explanation of our other finding, the relationship between IOS and team reasoning score, which is now conceived as cognitive hierarchy theory with a model of level-0 players who trade off label salience and payoff salience. The relationship between IOS and Team Reasoning Score also held at the level of individual questions that were designed to discriminate between those who picked the highest card and team reasoning, and for participants’ self-reported strategies, with those who reported

picking the odd-one-out (similar to label salience) having a higher IOS than those who reported picking the highest card (payoff salience).

Some other explanations of coordination from the psychology literature are not sufficient to explain our results. Psychologists have suggested that people may use evidential reasoning — whereby they expect others to behave in a similar way to themselves and therefore assume that, if they choose Hi, so will their partner — which is driven by the perception of similarity (Krueger, 2008). They refer to this as ‘social projection theory’. As a part of our priming tasks in Experiment 1, we told subjects in the *we* condition some ways in which their partner was similar to them and this was effective: participants in the *we* condition judged their partner as more similar to them than partners in the *they* condition. Nevertheless, there was no difference in behavior between the groups. Further, the correlation between IOS and team reasoning score changes little when perceived similarity is statistically controlled. Therefore, we conclude that our results cannot be explained solely by social projection theory. Other theories, such as Stackelberg reasoning (Colman et al., 2014; Colman et al., 2008; Pulford et al., 2014), which have been used to explain play in mixed motive games, do not seem to offer any plausible connection between IOS and behaviour, and therefore any explanation for the association of IOS and our Team Reasoning Strategy.

Our results suggest that a full theory of coordination will need to have an explanation for the relationship between social cohesion and coordination. We also found that friends play different strategies from strangers in coordination games. More generally, friends play differently from strangers and that may impact differently on different classes of games (Chierchia, Tufano & Coricelli, 2020). This finding has significance for the growing literature about group identification and social incentives as determinants of behaviour in the workplace (e.g., Akerlof & Kranton, 2005; Ashraf & Bandiera, 2018; Bandiera, Barankay & Rasul, 2010). It also has implications for the external validity of laboratory experiments. Economics experiments tend to make a virtue out of the anonymity of their participants, but most workplace behaviour occurs amongst colleagues who are not strangers. If play amongst those who know each other is different from those who are strangers, then researchers will need to take that into account.

## References

- Akerlof, G. A., & Kranton, R. E. (2005). Identity and the economics of organizations. *Journal of Economic Perspectives*, 19(1), 9–32.
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology*, 63(4), 596–612.
- Ashraf, N., & Bandiera, O. (2018). Social incentives in organizations. *Annual Review of Economics*, 10, 439–463.

- Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, *11*(2), 150–166.
- Attanasi, G., Hopfensitz, A., Lorini, E., & Moisan, F. (2016). Social connectedness improves co-ordination on individually costly, efficient outcomes. *European Economic Review*, *90*, 86–106.
- Bacharach, M. (2006). *Beyond individual choice: teams and frames in game theory*. Princeton: Princeton University Press.
- Bacharach, M., & Bernasconi, M. (1997). The variable frame theory of focal points: An experimental study. *Games and Economic Behavior*, *19*(1), 1–45.
- Bacharach, M., & Stahl, D. O. (2000). Variable-frame level-n theory. *Games and Economic Behavior*, *32*(2), 220–246.
- Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes & Intergroup Relations*, *12*(4), 533–547.
- Balliet, D., Wu, J., & De Dreu, C. K. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, *140*(6), 1556–81.
- Bandiera, O., Barankay, I., & Rasul, I. (2010). Social incentives in the workplace. *The Review of Economic Studies*, *77*(2), 417–458.
- Bardsley, N., Mehta, J., Starmer, C., & Sugden, R. (2010). Explaining focal points: cognitive hierarchy theory versus team reasoning. *The Economic Journal*, *120*(543), 40–79.
- Brewer, M. B., & Gardner, W. (1996). Who is this "we"? Levels of collective identity and self representations. *Journal of Personality and Social Psychology*, *71*(1), 83–93.
- Byrne, D. (1961). Interpersonal attraction as a function of affiliation need and attitude similarity. *Human Relations*, *14*(3), 283–289.
- Camerer, C. F., Ho, T. H., & Chong, J. K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, *119*(3), 861–898.
- Castro, M. F. (2008). Where are you from? Cultural differences in public good experiments. *The Journal of Socio-Economics*, *37*(6), 2319–2329.
- Cesario, J. (2014). Priming, replication, and the hardest science. *Perspectives on Psychological Science*, *9*(1), 40–48.
- Chen, R., & Chen, Y. (2011). The potential of social identity for equilibrium selection. *The American Economic Review*, *101*(6), 2562–2589.
- Chierchia, G., Tufano, F., & Coricelli, G. (2020). The differential impact of friendship on cooperative and competitive coordination. *Theory and Decision*, *89*(4), 423–452.
- Colman, A. M., & Gold, N. (2017). Team reasoning: Solving the puzzle of coordination. *Psychonomic Bulletin & Review*, 1–14.
- Colman, A. M., Pulford, B. D., & Lawrence, C. L. (2014). Explaining strategic coordination: Cognitive hierarchy theory, strong Stackelberg reasoning, and team reasoning. *Decision*, *1*(1), 35–38.

- Colman, A. M., Pulford, B. D., & Rose, J. (2008). Collective rationality in interactive decisions: Evidence for team reasoning. *Acta Psychologica, 128*(2), 387–397.
- Cookson, R. (2000). Framing effects in public goods experiments. *Experimental Economics, 3*(1), 55–79.
- Cornelissen, G., Dewitte, S., & Warlop, L. (2011). Are social value orientations expressed automatically? decision making in the dictator game. *Personality and Social Psychology Bulletin, 37*(8), 1080–1090.
- Crawford, V. P., Gneezy, U., & Rottenstreich, Y. (2008). The power of focal points is limited: even minute payoff asymmetry may yield large coordination failures. *The American Economic Review, 98*(4), 1443–1458.
- De Cremer, D., & Stouten, J. (2003). When do people find cooperation most justified? The effect of trust and self–other merging in social dilemmas. *Social Justice Research 16*(1), 41–52.
- Dienes, Z. (2016). How Bayes factors change scientific practice. *Journal of Mathematical Psychology, 72*, 78–89.
- Dorrough, A. R., & Glöckner, A. (2016). Multinational investigation of cross-societal cooperation. *Proceedings of the National Academy of Sciences, 113*(39), 10836–10841.
- Dorrough, A. R., Glöckner, A., Hellmann, D. M., & Ebert, I. (2015). The development of ingroup favoritism in repeated social dilemmas. *Frontiers in Psychology, 6*, 476.
- Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. Wiley. New York.
- Eckel, C. C., & Grossman, P. J. (2005). Managing diversity by creating team identity. *Journal of Economic Behavior & Organization, 58*(3), 371–392.
- Faillo, M., Smerilli, A., & Sugden, R. (2016). Can a single theory explain coordination? An experiment on alternative modes of reasoning and the conditions under which they are used. *CBES [Centre for Behavioural and Experimental Social Science] Working paper*, 16–01.
- Fallucchi, F., Luccasen, R. A., & Turocy, T. L. (2019). Identifying discrete behavioural types: a re-analysis of public goods game contributions by hierarchical clustering. *Journal of the Economic Science Association, 5*(2), 238–254.
- Fershtman, C., Gneezy, U., & Verboven, F. (2005). Discrimination and nepotism: The efficiency of the anonymity rule. *The Journal of Legal Studies, 34*(2), 371–396.
- Fiedler, S., Hellmann, D. M., Dorrough, A. R., & Glöckner, A. (2018). Cross-national in-group favoritism in prosocial behavior: Evidence from Latin and North America. *Judgment & Decision Making, 13*(1), 42–60.
- Gächter, S., Starmer, C., & Tufano, F. (2015). Measuring the closeness of relationships: a comprehensive evaluation of the 'Inclusion of the Other in the Self' scale. *PloS One, 10*(6), e0129478.
- Gächter, S., Starmer, C., & Tufano, F. (2017). Revealing the economic consequences of group cohesion. *IZA Discussion Paper*, 10824.
- Gold, N. (2012). Team reasoning, framing and cooperation. In Okasha, S., & Binmore, K. (Eds.). *Evolution and rationality: decision, cooperation and strategic behaviour*, (pp.

- 185–212). Cambridge: Cambridge University Press.
- Gold, N., & Colman, A. M. (2018). Team reasoning and the rational choice of payoff-dominant outcomes in games. *Topoi*, 1–12. <https://doi.org/10.1007/s11245--018--9575-z>.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26–34.
- Guala, F., Mittone, L., & Ploner, M. (2013). Group membership, team preferences, and expectations. *Journal of Economic Behavior & Organization*, 86, 183–190.
- Harris, C. R., Coburn, N., Rohrer, D., & Pashler, H. (2013). Two failures to replicate high-performance-goal priming effects. *PloS one*, 8(8), e72467.
- Hilbig, B. E., Glöckner, A., & Zettler, I. (2014). Personality and prosocial behavior: Linking basic traits and social value orientations. *Journal of Personality and Social Psychology*, 107(3), 529–539.
- Hilbig, B. E., Zettler, I., Leist, F., & Heydasch, T. (2013). It takes two: Honesty–Humility and Agreeableness differentially predict active versus reactive cooperation. *Personality and Individual Differences*, 54(5), 598–603.
- Isoni, A., Poulsen, A., Sugden, R., & Tsutsui, K. (2013). Focal points in tacit bargaining problems: Experimental evidence. *European Economic Review*, 59, 167–188.
- Jackson, J. W. (2008). Reactions to social dilemmas as a function of group identity, rational calculations, and social context. *Small Group Research*, 39(6), 673–705.
- Karpus, J. & Gold, N. (2016) Team reasoning: Theory and evidence. In J. Kiverstein (ed) *Handbook of philosophy of the social mind*, (pp. 400–417). New York: Routledge.
- Kagel, J. & McGee, P. (2014). Personality and cooperation in finitely repeated prisoner's dilemma games. *Economics Letters*, 124(2), 274–277.
- Kiyonari, T., & Yamagishi, T. (2004). Ingroup cooperation and the social exchange heuristic. *Contemporary Psychological Research on Social Dilemmas*, 269–286.
- Kramer, R. M., & Brewer, M. B. (1984). Effects of group identity on resource use in a simulated commons dilemma. *Journal of Personality and Social Psychology*, 46(5), 1044–57.
- Krueger, J. I. (2008). From social projection to social behaviour. *European Review of Social Psychology*, 18(1), 1–35.
- Mehta, J., Starmer, C., & Sugden, R. (1994). Focal points in pure coordination games: An experimental investigation. *Theory and Decision*, 36(2), 163–185.
- Pashler, H., Rohrer, D., & Harris, C. R. (2013). Can the goal of honesty be primed? *Journal of Experimental Social Psychology*, 49(6), 959–964.
- Perdue, C. W., Dovidio, J. F., Gurtman, M. B., & Tyler, R. B. (1990). Us and them: Social categorization and the process of intergroup bias. *Journal of Personality and Social Psychology*, 59(3), 475–86.
- Platow, M. J., Durante, M., Williams, N., Garrett, M., Walshe, J., Cincotta, S., Lianos, G., & Barutcu, A. (1999). The contribution of sport fan social identity to the production of prosocial behavior. *Group Dynamics: Theory, Research, and Practice*, 3(2), 161–169.

- Pulford, B. D., Colman, A. M., & Lawrence, C. L. (2014). Strong Stackelberg reasoning in symmetric games: An experimental replication and extension. *PeerJ*, 2, e263.
- Rohrer, D., Pashler, H., & Harris, C. R. (2015). Do subtle reminders of money change people's political views? *Journal of Experimental Psychology: General*, 144(4), e73.
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16(2), 225–237.
- Sally, D. (1995). Conversation and cooperation in social dilemmas. *Rationality & Society*, 7(1), 58–92.
- Schelling, T. C. (1960). *The strategy of conflict*. Cambridge, Mass., Harvard University Press.
- Stahl, D. O., & Wilson, P. W. (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization*, 25(3), 309–327.
- Sugden, R. (1993). Thinking as a team: Towards an explanation of nonselfish behavior. *Social Philosophy and Policy*, 10(01), 69–89.
- Sugden, R. (1995). A theory of focal points. *The Economic Journal*, 533–550.
- Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.
- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behaviour. In S. Worchel & W. G. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Chicago, IL: Nelson-Hall.
- Vadillo, M. A., Hardwicke, T. E., & Shanks, D. R. (2016). Selection bias, vote counting, and money-priming effects: A comment on Rohrer, Pashler, and Harris (2015) and Vohs (2015). *Journal of Experimental Psychology: General*, 145(5), 655–663.
- Volk, S., Thöni, C. & Ruigrok, W. (2011). Personality, personal values and cooperation preferences in public goods games: A longitudinal study. *Personality and Individual Differences*, 50(6), 810–815.
- Volk, S., Thöni, C. & Ruigrok, W. (2012). Temporal stability and psychological foundations of cooperation preferences." *Journal of Economic Behavior & Organization* 8(2), 664–676.
- Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van der Maas, H. L. (2011). Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). *Journal of Personality and Social Psychology*, 100(3), 426–432.
- Ward Jr, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*. 58(301), 236–244.
- Yamagishi, T., Makimura, Y., Foddy, M., Matsuda, M., Kiyonari, T., & Platow, M. J. (2005). Comparisons of Australians and Japanese on group-based cooperation. *Asian Journal of Social Psychology*, 8(2), 173–190.