CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Discussion Paper No. 2017-09

Simon Gächter, Chris Starmer and Fabio Tufano

Revealing the Economic Consequences of Group Cohesion

May 2017

CeDEx Discussion Paper Series
ISSN 1749 - 3293

CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

The Centre for Decision Research and Experimental Economics was founded in 2000, and is based in the School of Economics at the University of Nottingham.

The focus for the Centre is research into individual and strategic decision-making using a combination of theoretical and experimental methods. On the theory side, members of the Centre investigate individual choice under uncertainty, cooperative and non-cooperative game theory, as well as theories of psychology, bounded rationality and evolutionary game theory. Members of the Centre have applied experimental methods in the fields of public economics, individual choice under risk and uncertainty, strategic interaction, and the performance of auctions, markets and other economic institutions. Much of the Centre's research involves collaborative projects with researchers from other departments in the UK and overseas.

Please visit http://www.nottingham.ac.uk/cedex for more information about the Centre or contact

Suzanne Robey
Centre for Decision Research and Experimental Economics
School of Economics
University of Nottingham
University Park
Nottingham
NG7 2RD
Tel: +44 (0)115 95 14763
suzanne.robey@nottingham.ac.uk

The full list of CeDEx Discussion Papers is available at

http://www.nottingham.ac.uk/cedex/publications/discussion-papers/index.aspx

Revealing the Economic Consequences of Group Cohesion

By SIMON GÄCHTER, CHRIS STARMER AND FABIO TUFANO*

May 30, 2017 - Revised Version available 2019-16

We introduce the concept of "group cohesion" to capture the economic consequences of ubiquitous social relationships in group production. We measure group cohesion, adapting the "oneness scale" from psychology. A comprehensive program of new experiments reveals the considerable economic impact of cohesion: higher cohesion groups are significantly more likely to achieve Pareto-superior outcomes in classic weak-link coordination games. We show that effects of cohesion are economically large, robust, and portable. We identify social preferences as a primary mechanism explaining the effects of cohesion. Our results provide proof of concept for group cohesion as a productive new tool of economic research. (JEL C92, D03)

* Gächter: CeDEx, School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, UK (e-mail: simon.gaechter@nottingham.ac.uk). Starmer: CeDEx, School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, UK (e-mail: chris.starmer@nottingham.ac.uk). Tufano: CeDEx, School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, UK (e-mail: fabio.tufano@nottingham.ac.uk). We are grateful to Robert Cialdini and John K. Maner for sharing their experimental tools. We thank Antonio Alonso Aréchar, Michalis Drouvelis, Francesco Fallucchi, Natalia Montinari, Daniele Nosenzo, and Simone Quercia for assistance in running the experiments. We have benefited from numerous discussions with colleagues including Abigail Barr, Jordi Brandts, Colin Camerer, Robin Cubitt, Gary Charness, Yan Chen, David Cooper, Enrique Fatas, John Hillas, Muriel Niederle, John List, Carla Rampichini, Al Roth, Marie-Claire Villeval, Roberto Weber, and Frans van Winden. We also thank participants at various conferences, workshops and seminars for comments and suggestions. Research funding from the University of Nottingham is gratefully acknowledged. Tufano thanks also MIUR for financial support. Starmer and Gächter acknowledge the support of the ESRC funded Network for Integrated Behavioural Science (ES/K002201/1) and Gächter from the European Research Council grant ERC-AdG 295707 COOPERATION. The research reported in this paper received approval from the Research Ethics Committee of the School of Economics at the University of Nottingham. The authors declare no relevant or material financial interests that relate to the research described in this paper.

A vast array of economic and social activities occurs in groups. People need to coordinate and cooperate as colleagues in the workplace, teams on sports fields, army units on the battlefield, and across a host of less formal interactions with relatives, friends, and neighbors. While economic analysis has made great progress in understanding various incentive problems of teams and organizations (e.g., Gibbons and Roberts (2013)), little is known about how the psychological qualities of *social relationships* among group members influence group outcomes. In this paper, we demonstrate that this is an important gap by revealing *group cohesion* as a powerful factor of team production. We achieve this via a large-scale program of laboratory experiments involving more than 800 participants. We show that group cohesion can be reliably measured and that its effects on group performance are positive, large, robust and portable.

Members of *any* real human group inevitably have some form of relationship to other group members: for example, some people might feel "close" to other group members, whereas others may feel quite "distant." In our approach, we use the term group cohesion to refer to the state of the aggregate closeness ties within a group. The work reported here builds from our conjecture that group cohesion may be a significant factor of production influencing group performance, holding incentives and other factors constant.

Because relationship closeness is an essentially subjective concept, it is natural to wonder whether it can be reliably measured either for pairs of individuals, or aggregated to form a meaningful group-level statistic. Our research supports positive answers to both questions. Our measurement of group cohesion is based on the well-established "oneness scale" (Cialdini et al. (1997)) whose psychometric properties we replicated successfully in a companion paper (Gächter, Starmer and Tufano (2015)). The oneness scale asks subjects to indicate how close they feel to another person by using a simple pictorial tool due to Aron, Aron and Smollan (1992) and to state to what extent they would use the word "we" to characterize their relationship with that person. Both items are measured on a scale from 1 to 7 and the oneness measure is the average of responses on these two scales. From a measurement point of view, our innovation is to develop a simple aggregate statistic, based on oneness, to characterize the set of relationships within a group. Specifically, in experiments reported below, we ask each group member, privately, to indicate their oneness with every other group member. We then measure group cohesion as the group minimum of the individual average oneness ratings (details in Section II). As we show, this seemingly modest measurement innovation generates a tool with considerable explanatory power.

In a series of experiments, we demonstrate that group cohesion is an economically important production factor of group outcomes. A key feature of our setup is that we study the behavior of *real*, not artificially created, groups achieved by recruiting groups of friends to participate. We describe the details (including recruitment) in Section III and the characteristics of our real groups – which show wide variation in group cohesion – in Section IV.

Our workhorse to study group outcomes is a weak-link coordination game chosen because it captures economically interesting problems endemic to organizations and teams (e.g., Camerer and Weber (2013)). In our version of the weak-link game, inspired by Brandts and Cooper (2006), group members simultaneously choose an effort level. Payoffs to each group member then depend on their own effort and the lowest effort chosen by anyone (the "weakest link") in the group. The game has multiple strict Pareto-ranked Nash equilibria in material payoffs. This feature makes it particularly interesting for our purposes because it combines two dimensions of group success: features of coordination (choosing the same effort level as other group members) and cooperation (groups achieving Pareto-superior Nash equilibria). We expected our weak-link game to be a "harsh" environment in the sense that most groups who play this game under anonymity and in the absence of pre-existing social relationships will collapse to the Pareto-worst equilibrium and never escape from it (Brandts and Cooper (2006), and own replication).

As we show in Section V, group cohesion is a key determinant of behavior in our experiments: low cohesion groups usually descend rapidly to minimum effort; high cohesion groups fare much better and high cohesion appears necessary (though not sufficient) for achieving Pareto-superior outcomes. Surprisingly, our measure of group cohesion is the only variable that successfully predicts cooperation success; none of more than twenty control variables (demographics and group characteristics) explain minimum effort. Further experiments show that our results are robust to the timing of oneness measurement (before or after play of the weak-link game). By benchmarking our results against the effect of monetary incentives, we also show (Section VI) that the effort-enhancing effects of group cohesion are sizeable: large financial incentives are needed to achieve the levels of minimum effort expected for high cohesion groups.

In Section VII we turn to an explanation of our results. A rational choice perspective suggests three natural channels through which group cohesion could operate: it might affect some combination of group members' social preferences, their beliefs or the form of their strategic reasoning. Considering *social preferences*, it is plausible to assume that members of highly

cohesive groups care about one another and so place weight on each other's earnings.[1] In our weak-link game, if players do draw utility from each other's earnings, this is tantamount to (some) sharing of earnings, which reduces strategic risk and fosters coordination on Pareto-superior equilibria.[2] In relation to *beliefs*, highly cohesive groups may be more confident in simulating other group members' thought processes and likely actions, perhaps because of a history of interactions in different (related) situations, which allows for implicit learning (e.g., Holyoak and Spellman (1993), Rick and Weber (2010)). Finally, group cohesion might influence the nature of *strategic reasoning* in more substantive ways. For instance, according to one model of strategic thinking, "team reasoning" (e.g., Sugden (2003), Bacharach (2006)), people think in terms of what would be best for the team (e.g., picking the Pareto-best equilibrium) and are inclined to do their part in implementing the group-optimal outcome. An interesting possibility is that team reasoning may be more likely the more cohesive the team is. These three channels might operate jointly and potentially reinforce each other in high cohesion groups. By contrast, low cohesion groups may have low levels of social preferences, little implicit learning to draw on from shared situations, and no team perception to facilitate team reasoning.

We probe these possibilities in two steps. We first show that subjects who report high oneness with their fellow group members are indeed more likely to expose themselves to the strategic risk of choosing high initial effort in our weak-link games; they are also less "harsh" in their responses when others' effort levels are below their own. In highly cohesive groups, these tendencies apply across group members promoting coordination on equilibria above the Pareto-worst.

Our second step is to identify the social preferences channel as a promising route for explaining observed effects of group cohesion. We demonstrate this via additional experiments in which unrelated and anonymous group members play weak-link games but with *all earnings shared equally*. We interpret this manipulation as inducing a limiting form of social preferences (where all put equal weight on everyone's material payoffs). The results show patterns of effort (opening levels and dynamics) very comparable to the top third most cohesive groups from our main experiment. Thus, social preferences provide a parsimonious candidate explanation of how group cohesion promotes Pareto-superior equilibria.

[1] Recent work in anthropology (e.g., Hackman et al. (2017)) suggests a positive link between altruism and oneness.

[2] Chen and Chen (2011) develop a model in which social preferences reduce the strategic risk in coordination games. They also provide supporting experimental evidence from artificial groups with induced group identities).

I. Related Literature

Before presenting the substance of our paper we briefly place it in the literature. Our first contribution is to introduce the novel psychological concept of group cohesion. The concepts of relationship closeness and oneness are firmly established in the psychology literature (see next section) but barely considered in economics. The concept of group cohesion as derived from oneness is, to our knowledge, entirely new.[3] Moreover, while psychologists mostly use oneness to study intimate relationships, we extend its use to a broader variety of relationships providing economists with new tools to study them.

Our second contribution is to the experimental literature on coordination games, which hitherto has largely studied coordination among anonymous individuals without considering their social relationships. This research, summarized in Ochs (1995), Camerer (2003) and Devetag and Ortmann (2007), highlights the importance of structural features that facilitate coordination on efficient equilibria such as: communication (e.g., Cooper et al. (1992); Brandts and Cooper (2007)); leadership (Weber et al. (2001)); individual incentives (Brandts and Cooper (2006)); group size (Weber (2006)); choice of group members (Riedl, Rohde and Strobel (2016)); and culture at the organizational (Weber and Camerer (2003)) and even the societal level (Engelmann and Normann (2010)). Dispensing with anonymity, we show that the social psychological property of group cohesion is an independent and powerful production factor of group outcomes.

Our third contribution is to a growing literature on the economic nature of groups (e.g., Charness, Rigotti and Rustichini (2007); Hargreaves Heap and Zizzo (2009)), in particular the role of identity in organizations (e.g., Gibbons (2010); Akerlof and Kranton (2005); Bandiera, Barankay and Rasul (2010)) and the importance of understanding social-psychological dimensions of employment relationships more generally (e.g., Baron and Kreps (2013)). The papers most closely related to ours in this strand of literature are Charness, et al. (2007) and Chen and Chen (2011). Charness, et al. manipulate the saliency of group membership showing that it improves both cooperation in a prisoner's dilemma and coordination rates in a battle of the sexes game. Chen and Chen show, using artificial group identities, that an enhanced group identity can improve cooperation. In

[3] We are aware of a literature that uses the label *group cohesion* to refer to the "sum of forces acting on members to remain in a group" (Salas et al. (2015), p. 366). It measures group cohesion, so defined, on multi-item questionnaires. Our concept of group cohesion, and its measurement based on experimentally measured closeness relationships using the oneness scale (details in the next section) is quite distinct from this other literature.

contrast to both papers, we study real groups and demonstrate that naturally occurring levels of group cohesion, as measured by our new tool, enhance group outcomes.

II. Measuring Group Cohesion using the Oneness Scale

Our study involves the development and application of a new tool: a simple and portable measure of group cohesion designed to summarize the social relationships that exist between members of any group. To this end, we piggy-back on an established literature which has developed tools designed to measure the nature and strength of bilateral relationships between pairs of individuals. This literature demonstrates that important (i.e., predictive) features of – possibly complex – bilateral relationships can be summarized by simple measurement tools, which ask subjects to report how close they feel towards another target person. Our strategy builds on this literature and takes it a step further by assuming that important aspects of relationships that exist within groups can be summarized in terms of features of the set of pairwise closeness relationships in the relevant groups. Hence, on our measure a group will be more cohesive to the extent that its members, collectively, feel closer to one another. Since individual judgments of relationship closeness will be its building blocks, as a step towards group cohesion measurement, we now describe the rationale for, and key properties of, tools for measuring relationship closeness.

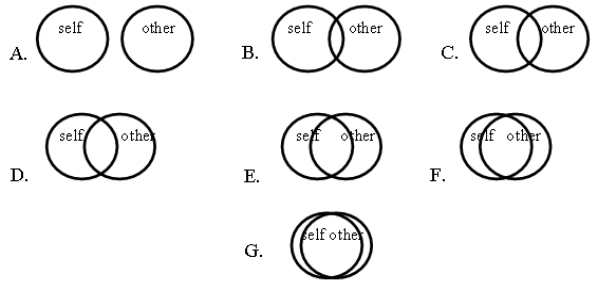
According to psychologists Kelley et al. (1983), relationship closeness depends on the frequency of interactions, the diversity of activities people undertake together, and the strength of influence people have on one another. In an effort to measure these determinants of relationship closeness, Berscheid, Snyder and Omoto (1989) developed a 69-item "Relationship Closeness Inventory." While this questionnaire is very detailed, it is not practical for many purposes. To provide a handy measurement technique, in a highly cited paper, Aron, et al. (1992) proposed a simple tool: the "Inclusion of the Other in the Self" (IOS) scale depicted in Fig. 1a. The IOS scale "is hypothesized to tap people's *sense* of being interconnected with another. That sense may arise from all sorts of processes, conscious or unconscious" (Aron, et al. (1992), p. 598). Essentially, it measures relationship closeness without examining its detailed determinants.

Aron, et al. argue that the IOS scale is highly intuitive, successful in tracking key dimensions of broader relationship closeness and has the decisive advantage of being very simple to implement. Subsequent research, most notably by Starzyk et al. (2006), developed an 18-item "Personal Acquaintance Measure" intended for application to a wider range of relationships including

acquaintances. Starzyk et al. found that their measure correlates strongly with the IOS scale. Together, these results make the IOS scale a very promising tool for our purposes.

Please, look at the circles diagram provided on your desk. Then, consider which of these pairs of circles best represents your connection with this person **before this experiment**. By selecting the appropriate letter below, please indicate to what extent **you and this person were connected**.

A. ☐ B. ☐ C. ☐ D. ☐ E. ☐ F. ☐ G. ☐



a. The "Inclusion of the Other in the Self" (IOS) scale

Please, select the appropriate number below to indicate to what extent, **before this experiment**, you would have used the term "**WE**" to characterize you and this person.

	1	2	3	4	5	6	7	
Not at all	☐	☐	☐	☐	☐	☐	☐	Very much so

b. The We Scale

FIGURE 1. ONENESS ELICITATION AS EXPLAINED TO THE PARTICIPANTS

To assess the robustness and reliability of the IOS scale, in a companion paper (Gächter, et al. (2015)) we replicated Berscheid, Snyder and Omato's Relationship Closeness Inventory, Starzyk et al.'s Personal Acquaintance Measure, and Aron, Aron and Smollan 's IOS scale, using 772 workers on Amazon Mechanical Turk as subjects. The results replicate remarkably closely the findings reported in these earlier studies and confirm the previously claimed psychometric properties of the Relationship Closeness Inventory, the Personal Acquaintance Measure, and the IOS scale. In Gächter, et al. we also performed a principal-component analysis of the questionnaire-based measures and found that it correlates at the 0.85 level with the IOS scale. We conclude from this that the IOS scale is a psychologically meaningful and reliable tool for measuring subjective closeness of relationships.

In our measurement of relationship closeness, we follow Cialdini, et al. (1997) who proposed to complement the IOS scale with the "We scale," depicted in Fig. 1b. Cialdini et al. call their measure – which is the average of the IOS scale and the We scale – the "oneness scale", and they interpret it as reflecting a "sense of shared, merged, or interconnected personal identities" (p. 483). In Gächter, et al. (2015) we confirmed Cialdini et al.'s claim that oneness correlates slightly better with the underlying survey instruments than the IOS scale alone and hence we use the oneness scale for our analysis.

In our experiments, we deploy the oneness scale as follows (wider procedural details are in Section III). Subjects participate as groups of four and each one rates three other identified group members separately and privately by responding to the IOS and We scales as depicted in Fig. 1. IOS scale responses are scored from A = 1 to G = 7. Oneness is the average of the IOS and We scale responses and therefore ranges from 1, "lowest oneness", to 7, "highest oneness."

We construct our group cohesion measure as follows. Since groups contain four people, each of whom gives a oneness rating for the three others in their group, each group generates twelve individual oneness ratings. Our group cohesion index is designed to capture some conception of the minimum oneness in a group; that is, to be a statistic for *how weak* is the weakest link in a group. We construct our measure in two steps. First, for each group member we select the minimum of the three oneness ratings they assign to others (i.e., their rating for the person they rate lowest). Group cohesion is then the average of these four numbers. We describe empirical properties of our measure, when applied to the real groups in our experiments, in Section IV.

III. Experimental Setup

A. *The Test Environment: The Weak-link Game*

The weak-link game has been widely studied in the lab, partly because it is often interpreted as representing a form of coordination problem endemic to organizations (e.g., Camerer and Weber (2013)). A classic example is workers preparing an aircraft for takeoff: the plane can only disembark once the *slowest* worker has fulfilled their task (Knez and Simester (2001)). Here, we describe the specific version of the game used in our main experiment and explain why it provides a good environment for studying possible effects of group cohesion.

We use a version of the weak-link game adapted from Brandts and Cooper (2006). A group of four players simultaneously choose one of five possible "effort levels" 1 to 5. The payoff to each player i is given by $\pi_i = 190 - 50e_i + 10b \cdot [\min(e_1, \ldots, e_4)]$ where e_i is player i's own effort, $\min(\cdot)$ is the lowest effort in the group, and b is a "bonus" rate controlling the marginal return to changes in minimum effort. In our main experiment, we set $b = 6$ mimicking Brandts and Cooper's baseline treatment. Table 1 illustrates the resulting payoff matrix in the format presented to subjects. Full instructions are available in Online Appendix A.

TABLE 1—THE PAYOFF MATRIX FOR THE WEAK-LINK GAME

		Minimum Effort				
		1	2	3	4	5
Effort by Player *i*	1	200				
	2	150	210			
	3	100	160	220		
	4	50	110	170	230	
	5	0	60	120	180	240

Note: The payoffs are expressed in points.

Each player chooses an effort level (i.e., a row of Table 1) and their payoff then depends on their own choice and the minimum effort among all members of their group (given by the column). From Table 1 the key tension embodied in the weak-link game is easy to see: everyone prefers that everyone chooses maximum effort (of 5) because this is the unique social optimum which simultaneously maximizes everyone's payoff (at 240 points); but the optimum may not be achieved because it is costly for any individual to exceed the minimum of efforts. On standard analysis, rational players will match their expectation of the minimum of others' efforts. The game has five strict Pareto-ranked equilibria on the diagonal of Table 1. Notice that the achievement of high payoffs requires elements of coordination (choosing the same effort level as other group members) and cooperation (groups achieving Pareto-superior Nash equilibria).

We chose this specification of a weak-link game for our main experiment for two primary reasons. First, Brandts and Cooper (2006) and Brandts and Cooper (2007) have established that this environment is "harsh" in the sense that, in the absence of aids to coordination (e.g., communication), behavior rapidly and reliably converges to the lowest Pareto-ranked equilibrium. Second, Brandts and Cooper also showed that people's behavior responds to financial incentives:

with b = 6 effort converges to the lowest possible level, but with higher bonuses, effort levels increase; a fact we exploit in a benchmarking exercise described in Section VI.

B. *Sampling Strategy and Sequence of Events*

The oneness measurement tool (Section II) and the harsh weak-link game just described lie at the heart of our experimental setup. Our strategy is as follows: we bring to the lab groups of subjects composed such that we expect variation in their group cohesion; we measure group cohesion and we observe behavior in a weak-link game. Hence, we can test for association between group cohesion and effort and, if we find it, we can explore its nature and power.

One important design challenge is how to create groups with suitable variation in pre-existing group cohesion. We constructed groups for what we refer to as our "main experiment" by sampling as follows. Invitations to prospective participants asked each invitee to bring three additional people who all knew each other and the invitee. Hence, participants (n = 260 students) arrived at the lab in sets of four acquaintances. Upon arrival, we assigned them to one of two matching protocols, the "Friends matching" (henceforth, F-matching; 47 groups) or the "Non-friends matching" (N-matching; 18 groups). In the F-matching, each quartet of acquaintances was allocated to the *same* group. By contrast, in the N-matching, each set of four acquaintances was split up so that each became a member of a *different* group. Thus, the only difference between the two matching protocols is that, under F-matching, group members are selected to have some prior history of social interactions with each other, whereas the N-matching aims to minimize the likelihood of prior social interaction, but keeping the recruitment procedures constant. Using these two matching protocols, we hoped to create variation in pre-existing cohesion across groups.

A second fundamental design issue is the potential for cross-task contamination. Specifically, given that we require participants to both provide oneness ratings of other group members and to play a (repeated) weak-link game, could the experience of one type of task affect behavior in the other? We addressed this issue in two ways. First, pilot experiments revealed that if measurement of oneness precedes the weak-link game, this does have some influence on minimum effort, compared to game play that is not preceded by the elicitation of oneness. A key question is then whether prior play of the game affects measured oneness. As a test of this, we ran within-subjects experiments (172 new participants; 43 groups, 27 F-matching, 16 N-matching) over two weeks. We refer to these as our "two-phase experiments". In week 1, we measured oneness and elicited

various individual-level characteristics. In week 2, the same subjects in the same groups played the weak-link game followed by elicitation of oneness ratings. Since it is plausible to assume that relationship closeness would not change systematically over the course of one week, any such changes in oneness ratings would be likely due to effects of the experience of game play.

Our results show that the oneness scores are not significantly different between week 1 and week 2 (individual average ratings as observations, Wilcoxon signed ranks test, $z = -1.033$, $p = 0.302$). This demonstrates an encouraging degree of test-retest reliability at the level of the individual. At the group level, the Spearman rank order correlation between week 1 and week 2 group cohesion is 0.928 ($n = 43$; $p < 0.001$). To further test the impact of game play on oneness ratings, we regressed changes in group cohesion on average minimum effort. The coefficient on minimum effort is insignificant (ordered probit, $\beta = -0.032$, $z = -0.28$, $p = 0.783$).[4] We conclude that prior play of the weak-link game has no detectable impact on subsequent measurement of oneness. This provides strong support for the sequence in our main experiment where we elicit the oneness ratings, for the construction of group cohesion, after the weak-link game.

C. *Procedures*

In all matching conditions, each group sat at a block of four computer workstations with partitions to prevent them from seeing each other's screens and responses. Each session started with an introduction read aloud by the experimenter. After that, each group of four participants was asked to stand up – one group at a time – so that each of its members could see the other members of their group. Subjects then followed computerized instructions, via their own screens. These first introduced the weak-link game followed by questions to test subjects' understanding of it. After the test, subjects played eight periods of the weak-link game. In each period, after each group member had (privately) entered their own effort level, their computer screen reported their own choice, their group's minimum, their own points for the current period, and their own accumulated points for all completed periods. Subjects knew that their total accumulated points across the eight periods would be converted to cash applying an exchange rate of 500 points = £1.00. For reasons explained above, oneness measurements were elicited after playing the game.

[4] We explored various other specifications involving the change in minimum effort between period 1 and 8; the initial minimum effort level; all effort levels; a variable representing the period (to capture a time trend) plus interactions between the period and effort levels. None of them revealed any systematic change in group cohesion in response to playing the weak-link game.

For this, after some computerized instructions, each participant was asked to focus on each other group member in turn and to respond, in sequence, to both the IOS scale and the We scale (Fig. 1) for the relevant focus person.

If group cohesion correlates with minimum effort, that could be either because it is a key explanatory factor or because it constitutes a good proxy to some subset of participants' characteristics that are, instead, the key factors determining effort. To address this, after the oneness ratings, participants completed a bank of "control tasks" designed to measure a range of individual-level characteristics from which we generate 25 individual-level variables. In addition to basic socio-demographic characteristics, these include indicators of attitudes to risk, group members or other people plus broader religious and political outlook. We report analysis of variance in group composition, based on the individual-level variables, in Section IV. We also construct group-level indicators which enter as controls in the analysis of the relationship between group cohesion and group minimum effort reported in Section V. Complete descriptions of the control tasks are in Online Appendix A; the individual-level variables and the construction of the group-level indicators are in the Online Appendix B.

We recruited participants via ORSEE (Greiner (2015)) and ran the experiments with z-Tree (Fischbacher (2007)) in the CeDEx lab at Nottingham University. Sessions lasted about one hour. Participants received task-related payoffs plus a £2.00 show-up fee. The mean payment was £7.88. Payments were made privately.

IV. The Environment of the Main Experiment: Cooperation Difficulty, Group Characteristics and the Distribution of Group Cohesion

Before presenting our primary results from the main experiment, we show that our experimental environment has three important characteristics that our setup was intended to induce. First, we aimed to create a "harsh" environment in the sense that groups with no significant history of prior social interaction would be characterized by severe cooperation failure (i.e., gravitating towards the lowest ranked equilibrium of the weak-link game). Using data from the N-matching, we find that, by period 8, almost 80 percent of groups collapse to minimum effort = 1; only two groups do

better achieving effort levels 2 and 3, respectively. Further tests reveal that this result is robust to game length and a non-student subject pool.[5]

A second key design objective was to import "real" groups into the lab. From the sociological literature on friendship we know that people tend to become friends with people who share similar socio-demographic backgrounds, attitudes and preferences (e.g., McPherson, Smith-Lovin and Cook (2001)); that is, "like-befriends-like." Hence, if we have succeeded in this second objective, for F-matching groups, we should find greater variation of socio-demographic and attitudinal characteristics *between* groups than *within* groups. We test this using the individual-level data generated from the control tasks.

Based on the non-parametric Kruskal-Wallis test, out of 25 characteristics the null of equal variance between and within F-matching groups is rejected at $p < 0.05$ for 17 and at $p < 0.10$ for 21 characteristics (more details are provided in Online Appendix B).[6] By contrast, and as expected, corresponding analysis for N-matching groups reveals no significant differences for any characteristic (5 percent level), comparing within and between group homogeneity.

A third, and crucially important, design objective was to create an environment in which there would be significant cross-group variation in group cohesion. Fig. 2, which plots the distribution of our measure of group cohesion, shows that we achieved this goal.

We note two encouraging features of this distribution. First, our recruitment and matching strategy has generated a wide range of group cohesion scores from a minimum of 1 to a maximum group cohesion score of 5.5. This wide range of group cohesion reflects a larger variation of oneness ratings between groups than within groups (Kruskal-Wallis test: χ^2 with ties = 145.4, $p < 0.001$). This means that our environment provides good scope for observing effects of variation in group cohesion, on group behavior, if such effects exist.

[5] To test whether a longer game might promote coordination on better equilibria, 32 fresh participants, recruited individually, played the game of Table 1 for 50 periods in fixed groups of four anonymous members. Six out of the eight groups were trapped into the lowest Pareto-ranked equilibrium by period 4; one by period 10; and one by period 22. We conclude that a long horizon does not help to escape inefficient equilibria. As a second robustness check, 44 fresh participants, recruited at the Joint Officer Training Program (JOTP) of the Swiss Army, played the game for 8 periods. These participants knew other group members were from the JOTP but not their exact identities. 73% of these groups collapsed to minimum effort from which we conclude that general knowledge about aspects of the identity of other group members is not enough to support cooperation.

[6] At $p < 0.05$ the significant variables are: Nationality, Age, Field of studies, Gender, Team perception, Religiousness, No. of club memberships, Prior group interactions, Share intention, Share influence, Monthly budget, No. of siblings, Cohabitee, City size, Trust index, Risk attitude, Self-financed. At $p < 0.10$ the significant variables are Political attitude, Empathy index, Income rank, Behavioral risk attitude. The insignificant variables are Trust strangers, Future happiness, Current happiness, Loss attitude.

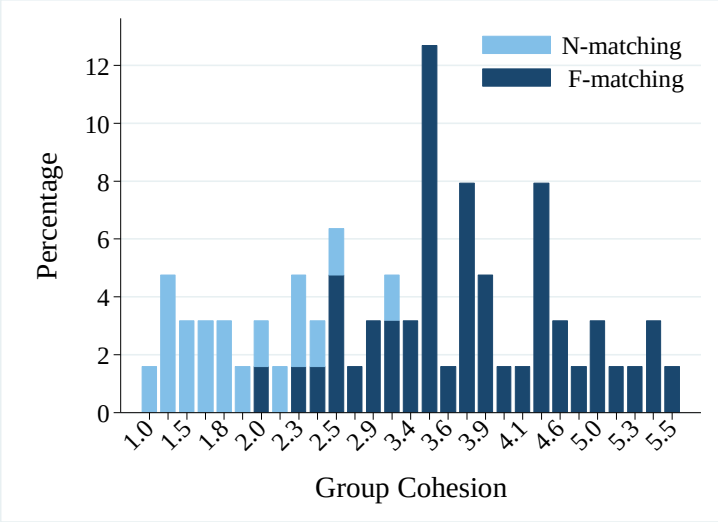


FIGURE 2. THE DISTRIBUTION OF GROUP COHESION UNDER F- AND N-MATCHING

Second, it is evident from Fig. 2 that group cohesion tends to be higher in the F-matching groups as compared to N-matching groups (means are 3.81 and 1.84, resp.; Mann-Whitney test, $z = 5.816$; $p < 0.001$). Since we have constructed F-matching groups to ensure that their members typically do have more pre-experimental friendship relationships than N-matching groups, this difference in the distributions of group cohesion across the matching protocols is what one should expect if oneness genuinely measures the closeness of social relationships.

V. The Relationships between Group Cohesion, Coordination and Cooperation

A. Descriptive Analysis

Fig. 3 presents scatter plots of minimum effort against group cohesion with separate panels for the first and last period of the weak-link game. Each plot also includes a line of best fit (OLS). We find a significant positive association between group cohesion and effort for both periods. Medium-to-high levels of group cohesion appear necessary for selecting high effort levels (i.e., minimum effort > 3). There is also evidence of some dynamic component revealed both by the change in concentration of observations across periods and picked up by the regression line which is both steeper and more strongly significant in period 8.

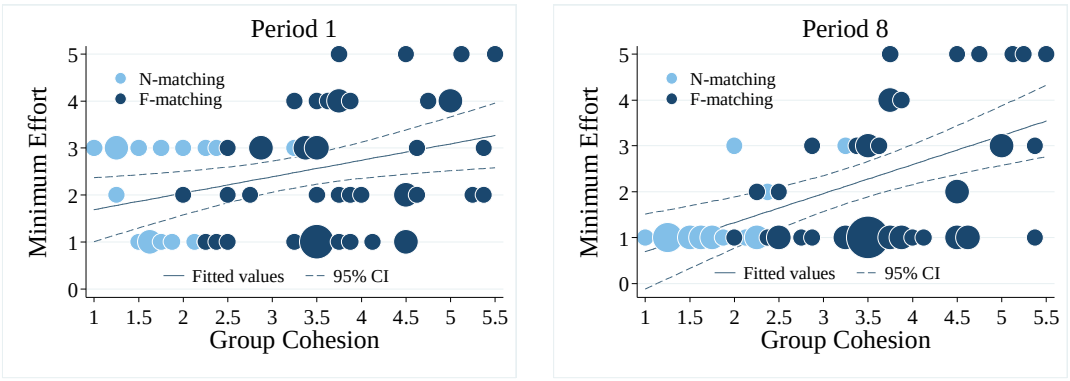


FIGURE 3. THE LINK BETWEEN GROUP COHESION AND GROUP-MINIMUM EFFORT

Notes: Size of symbols proportional to the number of observations. In Period 1, two N-matching observations are not displayed because they coincide with F-matching circles with coordinates (2.25, 1) and (2.5, 2); in Period 8, one N-matching observation is not displayed because it coincides with the F-matching circle at (2.5, 2). OLS Regression (65 groups), Period 1: $\beta = 0.313$ (se = 0.123, $p = 0.014$, $R^2 = 0.092$); Period 8 data: $\beta = 0.547$ (se = 0.123, $p < 0.001$, $R^2=0.240$). An ordered probit estimation generates qualitatively similar results.

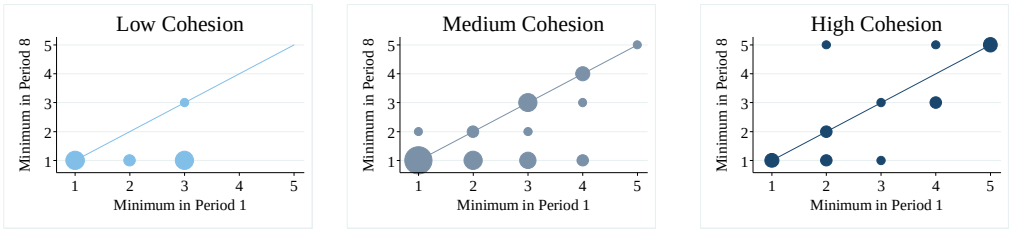
To further examine the dynamics suggested by Fig. 3, we separate the full set of 65 groups into three subsets of "low", "medium" and "high" cohesiveness groups (for details of partitions see note to Fig. 4). Fig. 4a illustrates how minimum effort changes between the first and last period and reveals marked differences in the patterns of transitions comparing across the three levels of cohesion. Notice that while almost all low cohesion groups collapse to minimum effort of 1, most high cohesion groups either preserve initial effort levels, or improve them.

Fig. 4b provides another window on these dynamics by showing the time path of (average) minimum effort, separately by partition. This reveals differences in both the initial levels of minimum effort and in the trends of minimum effort comparing across partitions: in contrast to low and medium cohesion groups, high cohesion groups cooperate more effectively in the initial period and do not experience a decay of minimum effort over time.

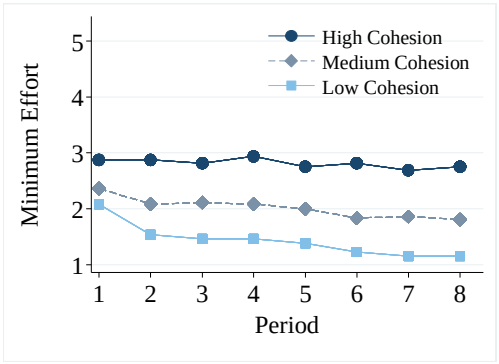
A final piece of descriptive analysis concerns "wasted effort" – the total of effort in a group above the group minimum in a particular period – which can be interpreted as a statistic for the extent of coordination failure. As Fig. 4c shows, average wasted effort in period 1 is around 5 and collapses to about 1 by period 8. Interestingly, cohesion levels do not substantially affect the dynamics of wasted effort.[7] The analyses of Figs. 4b and 4c suggest that group cohesion primarily facilitates cooperation (fostering decisions consistent with higher ranked equilibria), with

[7] We find only a weakly significant relationship between (average) group level wasted effort and group cohesion (Spearman's $\rho = -0.227$, $p = 0.069$; $n = 65$).

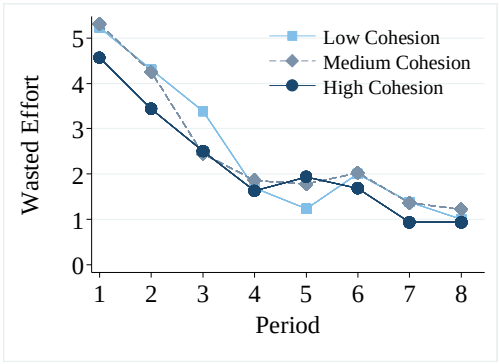
relatively little impact on coordination success (group members coordinating on the same equilibrium, regardless of its ranking). Indeed, the uniformly low rates of wasted effort by period 8 indicate strong convergence on equilibrium play for all levels of cohesion.



4a. Group minimum changes between period 1 and 8



4b. Group minima across periods



4c. Wasted effort

FIGURE 4. GROUP COHESION AND THE DYNAMICS OF COORDINATION

Note: "Low Cohesion" Partition (13 groups): *group cohesion* ∈ [1, 2]; "Medium Cohesion" Partition (36 groups): *group cohesion* ∈ (2, 4]; "High Cohesion" Partition (16 groups): *group cohesion* ∈ (4, 7]. Fig. 4a: size of symbol proportional to the number of observations. Fig. 4b: average group minimum effort over time. Fig. 4c: wasted effort per period, calculated as the sum of efforts in a group above the group minimum, averaged across groups.

B. The Predictive Power of Group Cohesion for Minimum Effort

The analysis just presented establishes a positive association between group cohesion and minimum effort. Yet, this correlation might be superficial if group cohesion is simply a proxy for other *observable* characteristics, which vary across groups. We now explore this issue via econometric analysis modelling minimum effort, across the 65 groups of the main experiment.

Table 2 summarizes the results from five different models. Apart from group cohesion and F-matching, all other variables are based on data from the control tasks and selected as follows.

TABLE 2—ORDERED PROBIT REGRESSIONS OF MINIMUM EFFORT ON GROUP COHESION AND BACKGROUND CHARACTERISTICS

Estimation method	GLLAM		Standard		Backw./Forw.
Controls for individual effects	Nested Random Effect		Clustering		Clustering
Dep. variable: Minimum Effort	*Model 1*	*Model 2*	*Model 3*	*Model 4*	*Model 5*
Group cohesion	2.506***		0.815***		0.448***
	(0.815)		(0.301)		(0.105)
F-matching	0.070	0.666	-0.122	0.032	
	(1.757)	(1.809)	(0.512)	(0.524)	
Gender	-1.645	-0.052	-0.514	-0.016	
	(2.638)	(2.587)	(0.997)	(1.001)	
Age	0.305*	0.110	0.087*	0.020	
	(0.160)	(0.151)	(0.049)	(0.044)	
Field of studies	-1.680	-1.188	0.519	-0.350	
	(2.294)	(2.333)	(0.798)	(0.804)	
Nationality	0.381	1.398	0.242	0.571	
	(3.293)	(3.200)	(1.144)	(1.160)	
No. of siblings	-0.453	-0.034	-0.165	-0.028	
	(0.754)	(0.743)	(0.317)	(0.312)	
City size	-1.537**	-1.266*	-0.467*	-0.332	
	(0.728)	(0.729)	(0.282)	(0.269)	
Cohabitee	0.498	0.493	0.143	0.148	
	(0.642)	(0.675)	(0.237)	(0.236)	
Monthly budget	-0.000	-0.001	-0.000	-0.001	
	(0.002)	(0.002)	(0.001)	(0.001)	
Self-financed	-0.015	-0.007	0.002	0.000	
	(0.021)	(0.022)	(0.008)	(0.008)	
No. of club memberships	0.041	-0.038	0.035	-0.061	
	(0.384)	(0.387)	(0.136)	(0.132)	
Prior group interactions	-0.194	-0.192	-0.044	-0.030	
	(0.483)	(0.509)	(0.150)	(0.157)	
Religiousness	0.100	-0.261	-0.007	-0.115	
	(0.383)	(0.372)	(0.155)	(0.153)	
Political attitude	-0.451	-0.132	-0.146	-0.054	
	(0.586)	(0.599)	(0.238)	(0.235)	
Share intention	0.837*	0.617	0.197	0.117	
	(0.499)	(0.499)	(0.171)	(0.177)	
Share influence	0.201	0.391	0.057	0.112	
	(0.473)	(0.485)	(0.172)	(0.170)	
Empathy index	-0.879	-0.238	-0.165	0.019	
	(0.692)	(0.689)	(0.277)	(0.258)	
Team perception index	-0.633	0.280	-0.194	0.107	
	(0.568)	(0.500)	(0.198)	(0.174)	
Trust index	-2.355	-2.243	-0.844	-0.761	
	(1.893)	(1.932)	(0.763)	(0.750)	
Risk attitude	0.513	0.570	0.178	0.184	
	(0.431)	(0.432)	(0.174)	(0.190)	
Log-(pseudo)likelihood	-3367.853	-372.349	-601.697	-626.084	-644.216
# level 1 (resp. 2) units	520 (65)	520 (65)	65	65	65

Notes: Data from Periods 1 to 8. Explanatory variables are at group level. Variable definition and construction are in Online Appendix B. Period dummies (always included, relative to Period 1) are significantly negative (at $p < 0.05$). Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

For each such variable, we re-ran the form of statistical analysis reported in Table B1 (comparing within and between group variance) but, in this case, we used the full data set including observations from both the F- and the N-matching protocols. Through this analysis, we identify the variables demonstrating *ex-ante* potential to explain variation in effort – that is, those showing

greater between (relative to within) group variation. We selected this subset of variables and constructed a group level statistic for each of them (see Online Appendix B for variable definition and construction). We included these group level statistics as controls for Model 1.

The specification of Model 2 is identical to that of Model 1 except that group cohesion is excluded. Since groups make multiple decisions, we need to account for interdependence of observations across periods. For Models 1 and 2, we do this by estimating the nested random model GLLAM (see Rabe-Hesketh, Skrondal and Pickles (2005)). Models 3 and 4 mirror, respectively, the specifications of Models 1 and 2. The only difference is the estimation procedure which, for Models 3 and 4, is a standard ordered probit with clustering at group-level.[8] Finally, Model 5 reports the results of a stepwise regression based on the same set of regressors as Models 1 and 3 with standard clustering (forward and backward selection produce identical results).

Considering first the models in which group cohesion is present (Models 1, 3 and 5), the results are striking: group cohesion emerges as the *only* variable with a consistently significant impact on minimum effort and its effects are always highly significant. Moreover, group cohesion is the unique survivor of the stepwise regression (Model 5).

The fact that very few other variables reach significance in Models 1, 3 and 5, could be due to them being highly correlated with group cohesion. But, while group cohesion does correlate with several background characteristics (see Online Appendix C), even when group cohesion is omitted (i.e., Models 2 and 4) no other variable achieves significance at a level better than 10 percent.

We conclude that group cohesion dominates our regression analysis as the uniquely consistent and highly significant predictor of minimum effort.

C. The Behavioral Impact of Group Cohesion

Having established a robust and highly significant association between group cohesion and minimum effort, we now examine the *magnitude* of this effect. We explore this using the conservative and parsimonious Model 5 (from Table 2) to predict the probabilities of each possible level of minimum effort, conditional on different levels of group cohesion.

[8] Using these two estimation procedures mirrors the strategy of Brandts and Cooper (2006). The rationale is to contain the true underlying model between two borders: the more powerful nested procedure is relatively prone to type I errors, while the standard clustering model is more prone to type II errors. Our main results, it turns out, are consistent across the two methods.

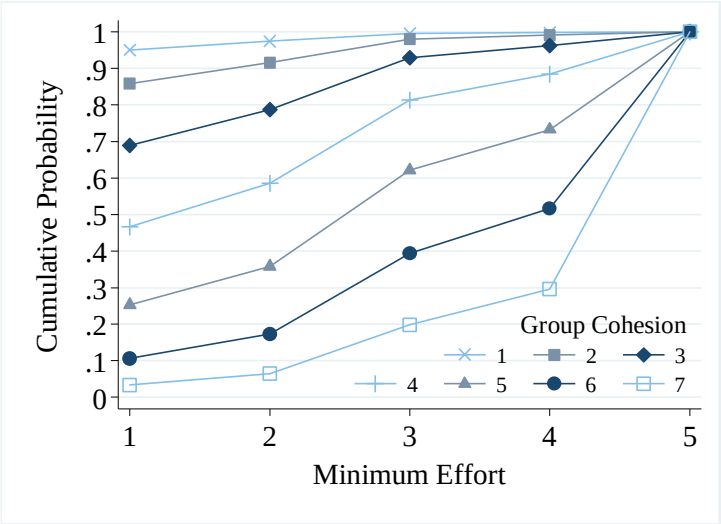


FIGURE 5. PREDICTED CDFS FOR MINIMUM EFFORT FOR EACH LEVEL OF GROUP COHESION (GROUP-LEVEL DATA FROM PERIOD 8)

Note: Post-estimation predictions obtained via SPost (Long and Freese (2005)) in STATA.

Fig. 5 presents the results for period 8 as cumulative distribution functions of minimum effort. The results demonstrate a sizeable predicted impact of group cohesion on minimum effort as we move between the extreme points of the group cohesion scale. For example, imagine a group characterized by minimum group cohesion (equal to 1). From the top line in Fig. 5 we see that such a group is almost certain to be at minimum effort (the actual probability of min. effort in this case is approximately 95 percent). By contrast, a group with maximum possible group cohesion (equal to 7) is very unlikely to end up at minimum effort 1 (probability of less than 4 percent) and is predicted to achieve minimum effort of at least 3 with a probability of about 80 percent.

D. Robustness and Portability

Two natural questions to ask about our results are: (i) how robust and replicable are they; and (ii) would they carry over to another context (beyond the weak-link game) where attainment of efficiency also requires cooperation? The data generated by our "two-phase" experiment (described in Section III.B) provides an opportunity to shed some initial light on both questions.

In relation to the first question we conduct a simple but informative check comparing average minimum effort across experiments (main vs two-phase experiment) using the partitions for group cohesiveness (i.e., low, medium, and high) introduced in Section V.A. These tests show that for both low and high cohesion groups, achieved minimum effort is the same in the main and the two-

phase experiment. By contrast, for groups with mid-range cohesion, minimum effort is *higher* for the two-phase experiment.

Why would it be that only the behavior of medium cohesion groups is affected by prior elicitation of oneness? We conjecture the following explanation. Recall that, in the two-phase experiment, we elicited the oneness ratings a week prior to the play of the weak-link game. The fact that this did not increase minimum effort for low or high cohesion groups is important because it demonstrates that the mere fact of eliciting oneness ratings does not act as a *general* prime that increases minimum effort; it is also consistent with the intuition that low cohesion groups would be aware that they have little relationship closeness; analogously, for high cohesion groups, eliciting oneness perhaps just confirms something that would already be clear to group members (their high relationship closeness). We speculate that mid-range cohesion groups, by contrast, may be less sure about the closeness of their relationships and eliciting it prior to game play, by virtue of focusing attention on it, may have enhanced their perceived cohesion and thereby minimum effort. (Additional supporting analysis can be found in Online Appendix D).

As a modest step in the exploration of portability, as part of the two-phase experiment, we also included a standard one-shot linear public good game (with groups of four and mpcr = 0.4) as the final task in the second week. For this task, each member of the group had to choose how many out of 20 tokens to contribute to a common pool (contributing 0 tokens is the money-maximizing choice and contributing 20 the socially efficient one). The data reveal a strongly significant positive association between group cohesion and total contributions to the public good.[9] This provides support for the expectation that cooperation enhancing effects of cohesiveness generalize beyond the weak-link game.

VI. The Economic Value of Group Cohesion

Fig. 5 shows that group cohesion has a strong behavioral effect on levels of cooperation. Here, we present a set of treatments designed to benchmark the cooperation enhancing effects of group cohesion against those of financial incentives. We present the results of four "bonus treatments" in which we varied the bonus (i.e., b in the payoff function π_i – see Section III.A) and hence the

[9] Correlation of group average contribution and group cohesion: Spearman's $\rho = 0.717$, $p < 0.001$, $n = 43$. Correlation of individual contribution to the public good and individual average oneness rating of the other group members: Spearman's $\rho = 0.567$, $p < 0.001$, $n = 172$.

payoff matrix of the weak-link games (between groups) holding everything else constant. The design of these treatments was extremely simple: in line with the earlier research by Brandts and Cooper and others, we recruited unrelated individuals (not groups of friends) and they completed 8 rounds of the weak-link game followed by a subset of the control tasks used in the main experiment (there was no elicitation of oneness ratings). The bonus rates in our four between-subjects treatments were set at 6, 14, 22 and 30, respectively. The first two bonus levels correspond with the lowest and highest bonus payments implemented by Brandts and Cooper (2006), while the other two go substantially higher in steps of 8 (the highest more than doubles their maximum). A total of 240 new subjects participated in the Bonus treatments (60 per treatment).

Fig. 6 summarizes the results by benchmarking the average of group minimum effort observed in each of the Bonus treatments against the group minimum efforts induced by group cohesion (the latter are predicted levels based on Model 5 in Table 2). In our data, increasing the bonus has a monotonic impact on expected minimum effort. At bonus level 6, average minimum effort is close to the minimum possible value of one and corresponds with the predicted minimum effort associated with low cohesion groups (i.e., a cohesion level of approximately 2). Increasing the bonus to the maximum used by Brandts and Cooper (i.e., b = 14) raises minimum effort to a level comparable to that expected from groups with a cohesion level of approximately 5.

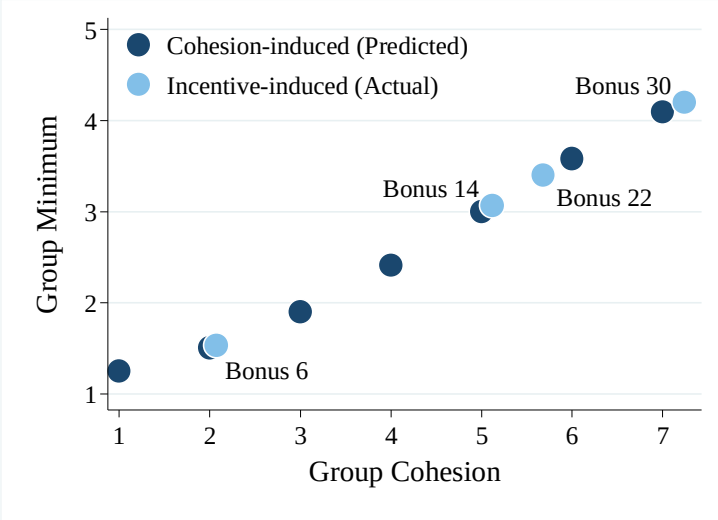


FIGURE 6. THE ECONOMIC VALUE OF GROUP COHESION

Perhaps the most striking feature of these data, however, is that to reach an expected minimum effort comparable to that associated with groups at maximum cohesion (i.e., group cohesion equal to 7) requires an almost five-fold increase in the baseline bonus rate to a value of around 30. These

results show that the economic value of group cohesion, measured by the material incentives needed to match its effort enhancing effects, is substantial.

VII. Towards an Explanation of the Power of Group Cohesion

In our final step, we move towards an explanation of how group cohesion fosters improvement in group outcomes. While our empirical strategy for measuring group cohesion has built on tools from social psychology, in seeking explanations for the effects of group cohesion our instincts as economists are to maintain two basic features of conventional economic analysis: group outcomes are to be modelled as a product of individual decisions, and individual decisions are to be understood through the lens of a rational choice model.

Within the rational choice framework, as noted in the introduction, there are three natural channels through which one might model the impact of group cohesion on individual decisions: via its impacts on preferences, beliefs, or reasoning processes. Existing work provides possible theoretical foundations focusing on different aspects of these channels. This includes models of group contingent social preferences (e.g., Chen and Chen (2011)); models of team reasoning (e.g., Sugden (1993), Sugden (1995) and Bacharach (1999), Bacharach (2006)) plus various theories of boundedly rational decision processes including cognitive hierarchy models (e.g., Stahl and Wilson (1995) and Camerer, Ho and Chong (2004)) and, most recently, the dual process model of Kets and Sandroni (2016).

Both theoretical and empirical considerations render it plausible to suppose that the different channels may be closely interconnected: they may operate in parallel and, potentially, reinforce each other. For example, from a theoretical point of view, if groups with higher cohesion care more about each other's payoffs, this reduces strategic risk, which in turn supports the expectation of higher effort levels within a group. Likewise, from an empirical point of view, since oneness is a function of the frequency and diversity of interactions (see Section II and Gächter, et al. (2015)), it is likely that high cohesion groups are ones in which members have had, correlated, opportunities to: form positive sentiments for each other (enhanced social preferences); have common experiences supporting implicit learning (enhanced beliefs); and develop some sense of team identity that helps promote cooperative reasoning ("I do my part in maximizing team success").

While this interconnectedness means that it will be difficult to separate cleanly the impacts of the different channels, in what follows, we demonstrate that the social preferences channel

provides a promising route for a parsimonious account of key patterns in our data. We make this case in two steps. In the first step, we highlight two behavioral tendencies of individuals in high cohesion groups. Specifically, relative to members of low cohesion groups, individuals in high cohesion groups deliver, on average, more effort in the first round; then, in subsequent rounds they also show "nicer" responses to effort which falls below their own. In a second step, we demonstrate that both of these behavioral tendencies can be replicated (qualitatively and quantitatively) in groups of *strangers* by experimentally *inducing a strong form of social preferences*.

We illustrate the two behavioral tendencies just mentioned via the two left hand panels of Fig. 7 showing the distribution of individual effort, by period, comparing groups with high (panel a) and low (panel b) group cohesion (these correspond with the two extreme partitions of Fig. 4). In these panels, average individual effort in each period is indicated with a circle.

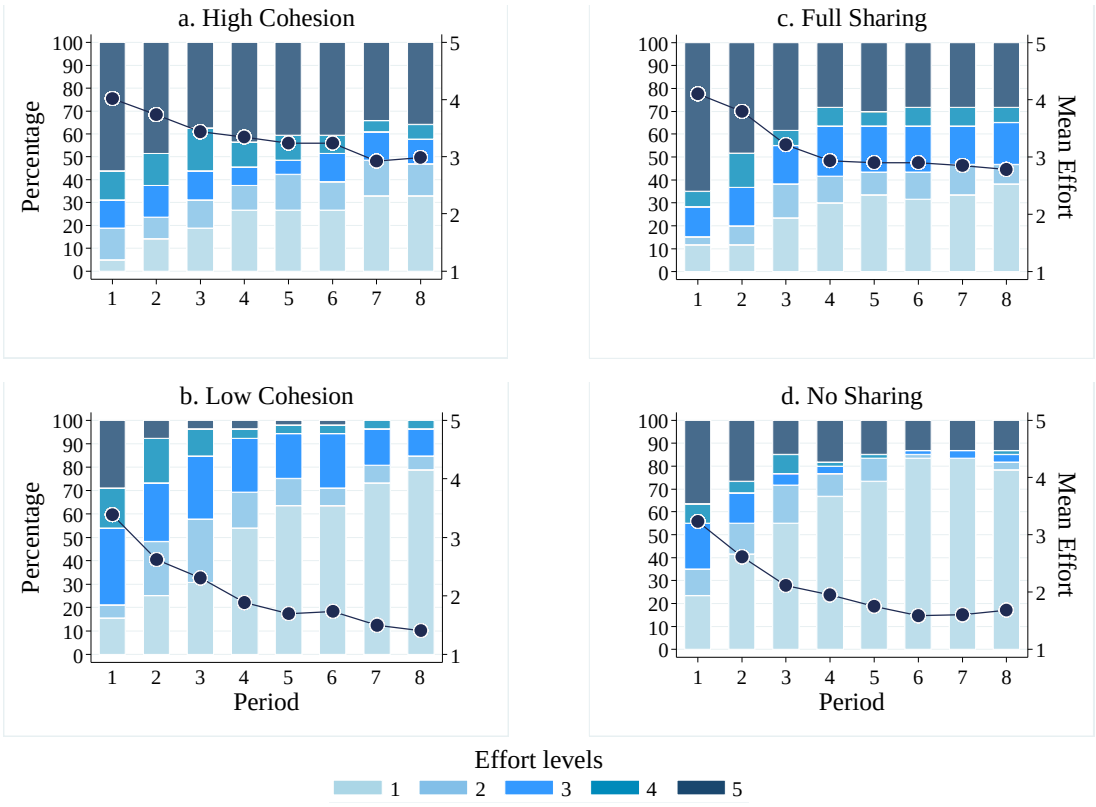


FIGURE 7. DISTRIBUTION OF INDIVIDUAL EFFORTS OVER TIME IN MAIN EXPERIMENT AND CONTROL EXPERIMENT WITH "INDUCED" SOCIAL PREFERENCES

Note: The "High Cohesion" Partition (panel *a*) has 16 groups from the main experiment; the "Low Cohesion" Partition (panel *b*) has 13 groups. Full Sharing (panel *c*) and No Sharing (panel *d*) have 15 groups each. The bars represent the percentage of each effort level ranging from 1 to 5. The y-axes show the relevant percentage. The connected dots represent mean efforts (individual level and measured on the secondary y-axes).

Notice that average effort in period 1 is higher for individuals in high cohesion groups: in period 1, 69 percent of effort choices in the high cohesion groups were at levels 4 or 5, whereas in the low cohesion groups only 46 percent were at that level. An econometric analysis also finds a highly significant positive influence of individual average oneness on individual effort choices.[10]

Comparison of the same pair of panels also reveals that average effort decays more slowly across periods for members of high cohesion groups (relative to low cohesion groups). By period 8, over 40 percent of individuals in high cohesion groups deliver effort above level 3, compared with less than 4 percent in low cohesion groups. We further examine these dynamics by focusing on each individual's change in effort following rounds in which they delivered above minimum effort. A subject who did not choose the minimum effort in period t is modelled as having a choice between three (mutually exclusive and exhaustive) options in period $t + 1$ which we label nice, moderate or harsh: *nice* agents deliver at least as much effort as before and maybe more; *moderate* agents reduce effort but no lower than the previous period minimum; *harsh* agents reduce their effort below the previous minimum. We conjectured that subjects with high average oneness ratings of their fellow group members would be more likely to be nice, while the reverse would be true for individuals with low average oneness ratings of their fellow group members. Econometric analysis strongly confirms this hypothesis.[11]

We suggest that any satisfactory explanation of the role played by group cohesion should be consistent with these patterns in the individual level data. We now demonstrate that the channel of social preferences provides a simple putative explanation of them. We achieve this via a new experiment involving 60 fresh participants recruited individually via standard procedures so that each subject typically did not know any other participant. In this experiment subjects played the weak-link game of Table 1 following the other procedures of the main experiment, but with the added feature that subjects were told we would *sum up all earnings and divide them equally among the four group member*s. We call this the *Full Sharing* treatment. With full sharing of payoffs, any subject can be expected to put equal weight on each group member's earnings including their own.

[10] GLAMM regression shows that individual effort tends to increase with the mean oneness rating of others in their group ($\beta = 0.139$; $p = 0.001$; main experiment data). Period dummies are negative ($p < 0.01$); the standard deviation of oneness ratings is not significant. Ordered probit analysis (clustered on individuals) confirms these conclusions.

[11] For observations where a subject did not choose the minimum effort in period t, an indicator variable for their choices in $t + 1$ assumes values of either 1, 0 or -1 for nice, moderate or harsh responses respectively. An ordered probit estimation (clustered on individuals; main experiment data) regressing the indicator variable on the individual mean of the three oneness ratings shows a significant and positive effect ($\beta = 0.101$, $p = 0.004$).

As such, one can interpret this treatment as implementing, in a very simple way, an extreme form of social preferences in which each agent places the same weight on the earnings of any group member (including themselves).[12]

We report the results for this treatment in Fig. 7c. For comparison, we benchmark against results from what we label here the *No Sharing* treatment in Fig. 7d (these are data from the *Bonus 6* experiment reported in Section VI). The results are striking. Notice that Full Sharing generates a pattern of individual efforts over time which is very similar to that observed for the high cohesion groups (Fig. 7a).[13] By comparison, No Sharing results in a pattern that is qualitatively comparable to the low cohesion groups from the main experiment.

These results demonstrate that inducing an extreme form of social preferences produces broad patterns in the individual effort decisions which closely mimic those of our high cohesion groups. While this analysis does not rule out other mechanisms also operating as drivers of the effort enhancing effects of group cohesion in our main experiment, it does highlight the channel of social preferences as a plausible and parsimonious candidate explanatory mechanism.[14]

VIII. Conclusions

It is hard to deny that social relationships affect many variables that naturally interest economists. More contentious is how much they matter and whether economic analysis could ever take account of them in a sufficiently parsimonious way to render the undertaking tractable and worthwhile. The research presented in this paper sheds new, and positive, light on these issues.

In this paper, we have explored the power of group cohesion – a hitherto unobservable characteristic and potential "production factor" of any real group – as a tool for predicting strategic behavior. Our previous related research has established that the oneness scale, on which our measurement of group cohesion is based, is simple to implement, highly portable and correlates

[12] Comparably extreme forms of control over beliefs or reasoning, supposing they were possible, would be of less interest. Consider controlling beliefs such that every subject believed every other subject would deliver maximum effort in the weak-link game. In this case, there is no social dilemma and it would be strange to observe anything other than convergence on the social optimum. A similar argument applies for the case of team reasoning. By contrast, while our manipulation of social preferences reduces strategic risk it does not eliminate that risk completely.

[13] Mann-Whitney tests for Fig. 7a vs. 7c: Period 1 ($z = -0.702$, $p = 0.483$); Period 8 ($z = 0.700$, $p = 0.484$). Fig. 7b vs. 7d: Period 1 ($z = 0.374$, $p = 0.7086$); Period 8 ($z = -0.403$, $p = 0.687$).

[14] Social preferences alone do not completely explain the effects of group cohesion. While individuals in the Full Sharing condition deliver comparable effort to those in high cohesion groups, those groups are not maximally cohesive (their average group cohesion is 4.83). Fig. 6 shows that, for maximally cohesive groups, the expected minimum effort is 4.10 which is considerably higher than the mean of minimum effort in our Full Sharing groups (2.60).

extremely well with more complex measures of personal relationships (Gächter, et al. (2015)). Using an extensive set of experiments involving over 800 subjects, and including a variety of robustness tests and benchmarking exercises, we have examined the predictive power of measured group cohesion in the context of experimental coordination games played by real groups which vary in the extent of pre-existing social relationships among their members.

Despite no possibilities for communication, in weak-link games, high cohesion groups do much better in terms of the equilibria they achieve and sufficiently high cohesion appears a necessary condition for cooperation. While group cohesion co-varies as expected with various objective characteristics of groups, surprisingly, in the presence of group cohesion no other factors appear to matter.

We have explored possible mechanisms that might underpin the impact of group cohesion and found evidence pointing tofor social preferences as a plausible candidate organizing individual-level behavioral tendencies.

We have shown that the cooperation enhancing effects of group cohesion are also large when benchmarked against the impact of financial incentives. While we cannot directly extrapolate to predict how big the effects of variation in group cohesion might prove to be in field contexts, our results provide strong support for exploring such questions using the oneness measurement tool.

Our results also have potential significance in the context of organisational performance (e.g., Akerlof and Kranton (2005); Bandiera, et al. (2010)). If group cohesion is a significant determinant of desirable team or group outcomes, then having a good tool to quantify it might facilitate a wide range of productive applied research. Finally, for those with interests in how to engineer better organisational or team performance, oneness measurement techniques may be valuable for assessing the impact of interventions, including the variety of team building-activities in which so many organisations already invest.

REFERENCES

Akerlof, George A., and Kranton, Rachel E. 2005. "Identity and the Economics of Organizations." *The Journal of Economic Perspectives*, 19(1): 9-32.

Aron, Arthur, Aron, Elaine N., and Smollan, Danny. 1992. "Inclusion of Other in the Self Scale and the Structure of Interpersonal Closeness." *Journal of Personality and Social Psychology*, 63(4): 596-612.

Bacharach, Michael. 1999. "Interactive Team Reasoning: A Contribution to the Theory of Co-Operation." *Research in Economics*, 53(2): 117-47.

Bacharach, Michael. 2006. *Beyond Individual Choice. Teams and Frames in Game Theory*. Princeton: Princeton University Press.

Bandiera, Oriana, Barankay, Iwan, and Rasul, Imran. 2010. "Social Incentives in the Workplace." *Review of Economic Studies*, 77(2): 417-58.

Baron, James N., and Kreps, David M. 2013. "Employment as an Economic and a Social Relationship." In *Handbook of Organizational Economics*, Robert Gibbons and John Roberts, 315-41. Princeton and Oxford: Princeton University Press.

Berscheid, Ellen, Snyder, Mark, and Omoto, Allen M. 1989. "The Relationship Closeness Inventory - Assessing the Closeness of Interpersonal Relations." *Journal of Personality and Social Psychology*, 57(5): 792-807.

Brandts, Jordi, and Cooper, David J. 2006. "A Change Would Do You Good ... An Experimental Study on How to Overcome Coordination Failure in Organizations." *American Economic Review*, 96(3): 669-93.

Brandts, Jordi, and Cooper, David J. 2007. "It's What You Say, Not What You Pay: An Experimental Study of Manager Employee Relationships in Overcoming Coordination Failure." *Journal of the European Economic Association*, 5(6): 1223-68.

Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton: Princeton University Press.

Camerer, Colin F. F., Ho, T. H., and Chong, J. K. 2004. "A Cognitive Hierarchy Model of Games." *Quarterly Journal of Economics*, 119(3): 861-98.

Camerer, Colin F., and Weber, Roberto A. 2013. "Experimental Organizational Economics." In *Handbook of Organizational Economics*, Robert Gibbons and John Roberts, Princeton and Oxford: Princeton University Press.

Charness, Gary, Rigotti, Luca, and Rustichini, Aldo. 2007. "Individual Behavior and Group Membership." *American Economic Review*, 97(4): 1340-52.

Chen, Roy, and Chen, Yan. 2011. "The Potential of Social Identity for Equilibrium Selection." *American Economic Review*, 101(6): 2562-89.

Cialdini, Robert B., Brown, Stephanie L., Lewis, Brian P., Luce, Carol, and Neuberg, Steven L. 1997. "Reinterpreting the Empathy-Altruism Relationship: When One into One Equals Oneness." *Journal of Personality and Social Psychology*, 73(3): 481-94.

Cooper, Russell, DeJong, Douglas V., Forsythe, Robert, and Ross, Thomas W. 1992. "Communication in Coordination Games." *The Quarterly Journal of Economics*, 107(2): 739-71.

Devetag, Giovanna, and Ortmann, Andreas. 2007. "When and Why? A Critical Survey on Coordination Failure in the Laboratory." *Experimental Economics*, 10(3): 331-44.

Engelmann, Dirk, and Normann, Hans-Theo. 2010. "Maximum Effort in the Minimum-Effort Game." *Experimental Economics*, 13(3): 249-59.

Fischbacher, Urs. 2007. "Z-Tree: Zurich Toolbox for Readymade Economic Experiments." *Experimental Economics*, 10(2): 171-78.

Gächter, Simon, Starmer, Chris, and Tufano, Fabio. 2015. "Measuring the Closeness of Relationships: A Comprehensive Evaluation of the 'Inclusion of the Other in the Self' Scale." *Plos One*, 10(6): e0129478.

Gibbons, Robert. 2010. "Inside Organizations: Pricing, Politics, and Path Dependence." *Annual Review of Economics*, 2(1): 337-65.

Gibbons, Robert, and Roberts, John eds**.** 2013. *The Handbook of Organizational Economics*. Princeton and Oxford: Princeton University Press.

Greiner, Ben. 2015. "Subject Pool Recruitment Procedures: Organizing Experiments with Orsee." *Journal of the Economic Science Association*, 1(1): 114-25.

Hackman, Joseph, Munira, Shirajum, Jasmin, Khaleda, and Hruschka, Daniel. 2017. "Revisiting Psychological Mechanisms in the Anthropology of Altruism." *Human Nature*, 28(1): 76-91.

Hargreaves Heap, Shaun P., and Zizzo, Daniel J. 2009. "The Value of Groups." *American Economic Review*, 99(1): 295-323.

Holyoak, Keith J., and Spellman, Barbara A. 1993. "Thinking." *Annual Review of Psychology*, 44(1): 265-315.

Kelley, Harold H. , Berscheid, Ellen, Christensen, Andrew, Harvey, John H., Huston, Ted L., Levinger, George, McClintock, Evie, Peplau, Letitia A., and Peterson, Donald R. 1983. *Close Relationships*. New York: Freeman.

Kets, Willemien, and Sandroni, Alvaro. 2016. "A Belief-Based Theory of Homophily." mimeo, Kellogg School of Management, Northwestern University.

Knez, Marc, and Simester, Duncan. 2001. "Firm-Wide Incentives and Mutual Monitoring at Continental Airlines." *Journal of Labor Economics*, 19(4): 743-72.

Long, Scott J., and Freese, Jeremy. 2005. *Regression Models for Categorical Outcomes Using Stata (2nd Ed.)*. College Station (TX): Stata Press.

McPherson, Miller, Smith-Lovin, Lynn, and Cook, James M. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology*, 27: 415-44.

Ochs, Jack. 1995. "Coordination Problems." In *Handbook of Experimental Economics*, John H. Kagel and Alvin E. Roth, 195-252. Princeton: Princeton University Press.

Rabe-Hesketh, Sophia, Skrondal, Anders, and Pickles, Andrew. 2005. "Maximum Likelihood Estimation of Limited and Discrete Dependent Variable Models with Nested Random Effects." *Journal of Econometrics*, 128(2): 301-23.

Rick, Scott, and Weber, Roberto A. 2010. "Meaningful Learning and Transfer of Learning in Games Played Repeatedly without Feedback." *Games and Economic Behavior*, 68(2): 716-30.

Riedl, Arno, Rohde, Ingrid M. T., and Strobel, Martin. 2016. "Efficient Coordination in Weakest-Link Games." *The Review of Economic Studies*, 83(2): 737-67.

Salas, Eduardo, Grossman, Rebecca, Hughes, Ashley M., and Coultas, Chris W. 2015. "Measuring Team Cohesion: Observations from the Science." *Human Factors*, 57(3): 365-74.

Stahl, Dale O., and Wilson, Paul W. 1995. "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior*, 10(1): 218-54.

Starzyk, K. B., Holden, R. R., Fabrigar, L. R., and MacDonald, T. K. 2006. "The Personal Acquaintance Measure: A Tool for Appraising One's Acquaintance with Any Person." *Journal of Personality and Social Psychology*, 90(5): 833-47.

Sugden, Robert. 1993. "Thinking as a Team: Towards an Explanation of Nonselfish Behavior." *Social Philosophy and Policy*, 10(01): 69-89.

Sugden, Robert. 1995. "A Theory of Focal Points." *The Economic Journal*, 105(430): 533-50.

Sugden, Robert. 2003. "The Logic of Team Reasoning." *Philosophical Explorations*, 6(3): 165-81.

Weber, R. A., and Camerer, Colin F. F. 2003. "Cultural Conflict and Merger Failure: An Experimental Approach." *Management Science*, 49(4): 400-15.

Weber, Roberto A. 2006. "Managing Growth to Achieve Efficient Coordination in Large Groups." *The American Economic Review*, 96(1): 114-26.

Weber, Roberto A., Rottenstreich, Yuval, Camerer, Colin, and Knez, Marc. 2001. "The Illusion of Leadership: Misattribution of Cause in Coordination Games." *Organization Science*, 12(5): 582-98.

Online Appendix for

Revealing the Economic Consequences of Group Cohesion

By SIMON GÄCHTER, CHRIS STARMER AND FABIO TUFANO

May 26, 2017

Table of Contents

Appendix A: Experimental Procedures and Instructions

Section A.I reports the invitation letter for the main experiment. Section A.II documents the wording used for the welcome of experimental participants and the initial oral instructions. Section A.III presents the experimental computerized instructions (note that only the key screenshots are shown – omitted screenshots can be visualized by running the z-Tree code provided as part of this online

supplement). Section A.IV details the post-experimental control tasks (including related summary statistics).

A.I. Invitation Letter

Dear #Name# #Surname#,
You registered with CeDEx to participate in experiments. We would like to invite YOU plus THREE of your friends to take part in our coming experiment.

IMPORTANT INFORMATION

To participate in this experiment, YOU are required to bring THREE people who know you as well as each other along with you.
Everyone among them needs to
1. …be someone you know (e.g. a friend, a housemate, someone on your course etc.).
2. …be someone who knows the other two people (e.g. a friend, a housemate, someone on your course etc.).
3. …be able to attend the session you have signed up for.

The experiment will not take more than 90 minutes, and at the end YOU and your THREE friends will be paid in private and in cash. We expect an average earning of £ 8 per participant.
The experiment will take place in the CeDEx computer lab, room C41, Sir Clive Granger Building. We are planning the following sessions (everyone can participate in one session only):
#Session list#
If YOU plus THREE of your friends would like to participate, please click on the link below to sign up to the session of your choice. After signing up, reply to this e-mail to communicate the full name of your THREE friends who will participate (Your friends do not have to sign up, but you must ensure that they attend the session with you).
Please note: people that sign up to a session and do not turn up cause us problems; sign up to a session only if you are sure you can attend it and that YOU are able to bring THREE friends as well. If you sign up please make sure YOU and your THREE friends do attend.
#Link#
(If the link does not work, copy it and paste it into the address field of your internet browser.)
This experiment is named FT1. If you email CeDEx regarding this experiment please include “FT1” in the subject field of your email to ensure that your message is dealt with promptly.

Best regards,
Fabio Tufano

A.II. Welcome and Oral Instructions

Welcome to the CeDEx Lab!
Thank you very much for participating.
This is an experiment in decision making and it is funded by CeDEx. The instructions are simple. If you follow them carefully you might earn an appreciable amount of money. These earnings will be paid to you privately, in cash, at the end of the experiment.
The experiment consists of several parts. You will learn about them as we go along. During the experiment, you will be required to make some choices and answer some questions. Note that all your responses and data entries will be kept anonymous.
We are interested in your individual choices. Therefore, communication is not allowed during the experiment. If you have a question, please just raise your hand.
In the experiment you will be a member of a group of four people. The groups will stay the same throughout the experiment. Shortly, in order to see who is in your group you will be asked by me to stand up in turn, group by group. When this happens, please pay attention to the composition of your group.
Please look at the computer monitor. On the top-left corner there is a sticker with a number. That is your participation number.
Participants from number 1 to 4, please stand up… Participants from number 5 to 8, please stand up… Participants from number 9 to 12, please stand up… Participants from number 13 to 16, please stand up…
(Participants from number 17 to 20, please stand up…)
Are there any questions?

A.III. Computerized Instructions: Script and Screenshots

Experimental instructions were only delivered on computer screens. In this subsection, we report the full experimental script in a series of block quotations marked by (*) with examples of key screenshots. Note that the text in italics reported in squared brackets was not seen by experimental participants.

The first computer screen presented to participants called their attention to Task 1. The text read as follows:

****

Task 1

In this task you will be asked to make several decisions which will affect your earnings from the experiment.
Continue [*Button*]

****

Participants were then introduced to the weak-link game with the following written instructions:

****

Instructions for Task 1

We are now going to ask you to make a series of decisions. There will be eight rounds. In every round, each member of your group including you will choose a value of X. The values of X you may choose are 1, 2, 3, 4, 5. The value you choose for X and the smallest value of X chosen by a participant in your group will determine the payoff you receive for that round.

In the box displayed to the right [*below in this appendix*] of this screen, you are provided with an EARNINGS TABLE that shows the potential payoffs that you may receive. Please look at this table now. Note that the EARNINGS TABLE is the same for every participant in the experiment.

Earnings are in points. The exchange rate is 500 points = 1 Pound. Only your earnings are shown in the EARNINGS TABLE. The earnings in each round may be found by looking across from the value you choose on the left-hand-side of the table and down from the smallest value of X chosen by a participant in your group.

For example, suppose the other members of your group choose 4, 5 and 4 respectively while you choose 3. Thus, the smallest value chosen is 3 and you earn 220 (in points) for that round. As a second example, if you choose 2 and the smallest value chosen is 1, you earn 150 for that round and so forth.

Note that you will be paid the total earnings from the eight rounds.

To be sure that everyone understands the instructions so far, please fill in the QUESTIONS on the next screen. When you are done, confirm your answers. If there are any mistakes in your answers, you will be asked further questions till you will get them right.

Continue [*Button*]

EARNINGS TABLE

		The smallest value of X chosen				
		1	2	3	4	5
Your Choice	1	200				
	2	150	210			
	3	100	160	220		
	4	50	110	170	230	
	5	0	60	120	180	240

****

Figure A1 below presents the screenshot of the "Instructions for Task 1."

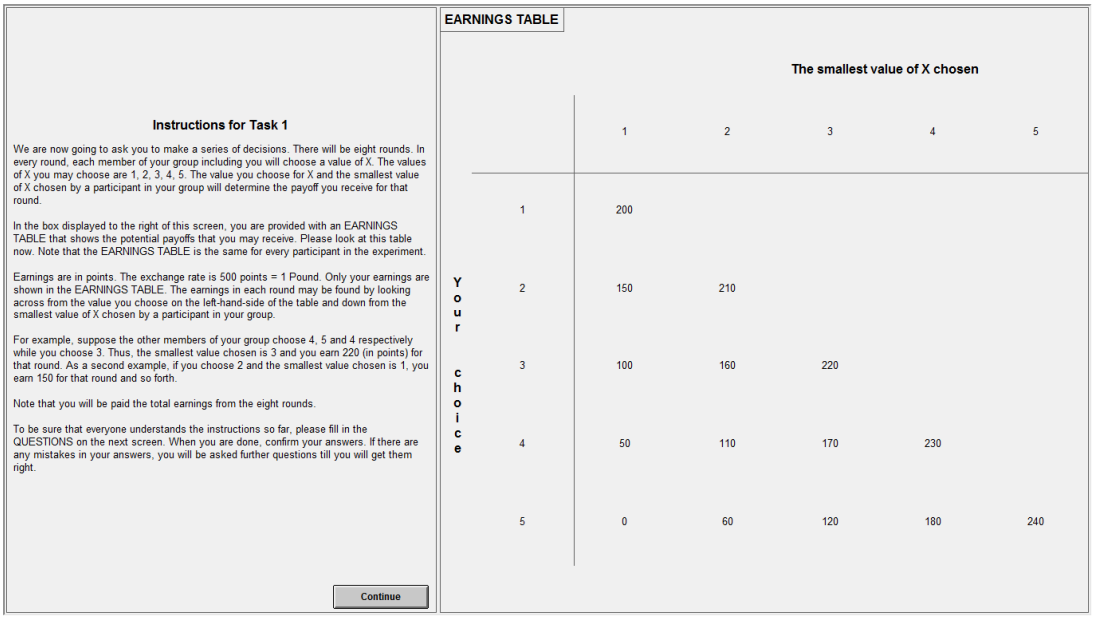


FIGURE A1. SCREENSHOT OF "INSTRUCTION FOR TASK 1"

The weak-link game instructions were followed by questions to test subjects' understanding of the game. The text of the first set of test questions is reproduced below:

****

If your choice of X is 2 and the smallest value of X chosen is 2, your points are [*Entry field*]
If your choice of X is 5 and the smallest value of X chosen is 4, your points are [*Entry field*]
If your choice of X is 3 and the smallest value of X chosen is 1, your points are [*Entry field*]
If your choice of X is 4 and the smallest value of X chosen is 4, your points are [*Entry field*]
Confirm [*Button*]

EARNINGS TABLE

The smallest value of X chosen

		1	2	3	4	5
Your Choice	1	200				
	2	150	210			
	3	100	160	220		
	4	50	110	170	230	
	5	0	60	120	180	240

****

After the test questions, participants played eight periods (the experimental instructions use the term "round") of the weak-link game. The relevant text for the first period of the game is reported as follows:

****

Round 1 out of 8
Please, enter the choice that you wish to make.
Your choice is [*Entry field*]
Press OK to confirm **Your choice**
OK [*Button*]

EARNINGS TABLE

		The smallest value of X chosen				
		1	2	3	4	5
Your Choice	1	200				
	2	150	210			
	3	100	160	220		
	4	50	110	170	230	
	5	0	60	120	180	240

****

Figure A2 below is the screenshot of the input screen for the first period of the weak-link game.

FIGURE A2. PERIOD 1 OF THE WEAK-LINK GAME, INPUT SCREEN

In each period, after each participant had entered their chosen effort their computer screen would report their period feedback consisting of their own effort, their group minimum effort, their points for

the current period and their total accumulated points for all completed periods. The script for Period 1 feedback is shown below:

> **********
>
> Round 1 out of 8
>
> In the previous period **Your choice** was and the minimum value was [*Output field*]
>
> Hence, your current payoff is [*Output field*] points while your points so far are [*Output field*]
>
> Continue [*Button*]
>
> **********

Once participants had completed eight periods of the weak-link game, they moved to Task 2. This was introduced with the following text:

> **********
>
> **Task 2**
>
> In this task you will be asked to focus your attention on [*Person X – Output field*] and answer a questionnaire.
> Continue [*Button*]
>
> **********

After preliminary instructions, participants were asked to rate an identified focus person on the We scale as follows:

> **********
>
> Please, focus your attention on [*Person X – Output field*].
> Note that in order to help you in focusing on the right person we have provided you with the "*Group layout*" below, which describes your positions in the block of Lab workstations.

Group Layout

Person X	
Person Y **You**	

> Please select the appropriate number below to indicate to what extent, **before this experiment**, you would have used the term "**WE**" to characterize you and this person.

	o	o	o	o	o	o	o	
Not at all	1	2	3	4	5	6	7	Very much so

Confirm [*Button*]

****

Then, subjects were provided with instructions and asked to rate the focus person on the IOS scale as described below:

****

Please, focus your attention on [*Person X – Output field*].
Note that in order to help you in focusing on the right person we have provided you with the "*Group layout*" below, which describes your positions in the block of Lab workstations.

Group Layout

Person X	
Person Y **You**	

Please, look at the circles diagram provided on your desk.
Then, consider which of these pairs of circles best represents your connection with this person **before this experiment**.
By selecting the appropriate letter below, please indicate to what extent **you and this person were connected**.

- o A.
- o B.
- o C.
- o D.
- o E.
- o F.
- o G.

Confirm [*Button*]

****

Figure A3 below reproduces the screenshot eliciting the IOS rating.

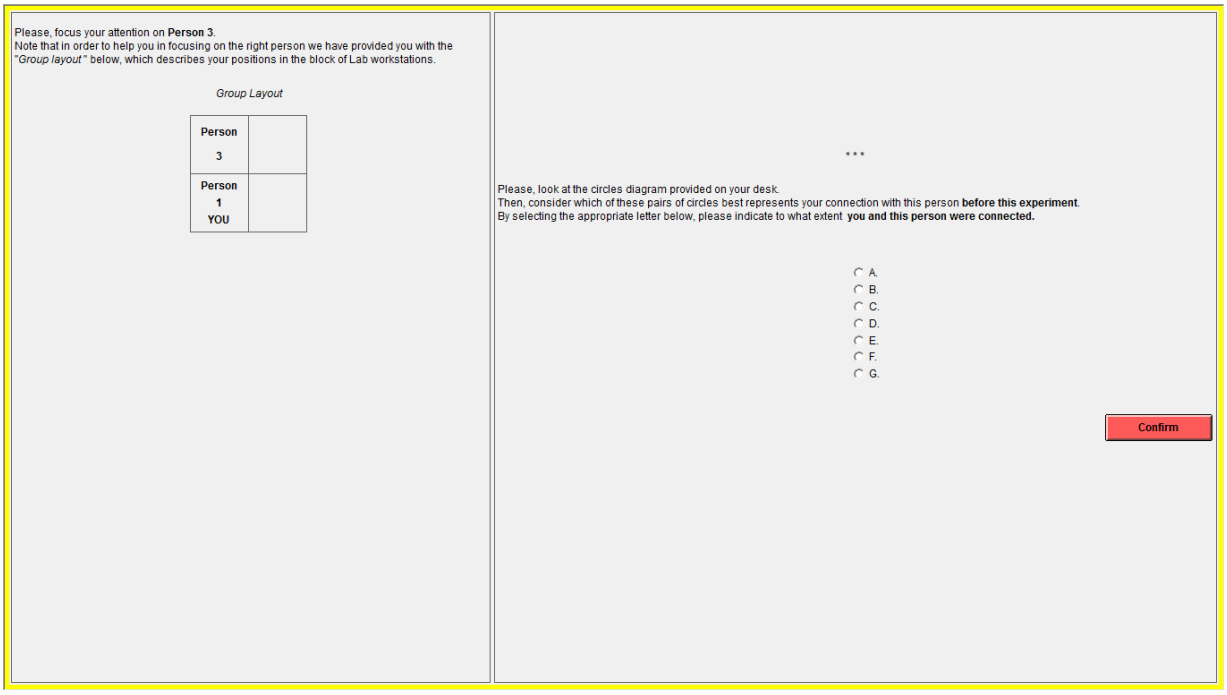


FIGURE A3. IOS SCALE, INPUT SCREEN

Note: Rater: Person 1; Focus Person: Person 3. The group was allocated to a block of four workstations with a square layout.

The elicitation of the We scale and the IOS measures was followed by a scenario description and a set of questions aimed at eliciting participants' willingness to help. The text read as follows:

****

Please, focus your attention on [*Person X – Output field*].
Note that in order to help you in focusing on the right person we have provided you with the "*Group layout*" below, which describes your positions in the block of Lab workstations.

Group Layout

Person X	
Person Y You	

Please generate an image in as much detail as you can, including this person's physical appearance, behaviors, and personality characteristics. Even if you are not sure about some aspects, we would like you to take your best guess. Then, keeping this person's image in mind we would like you to respond to questions about this person in the situation described in the scenario below. We are interested in how people react to a given situation in which another person needs their help. So for the scenario below, we would like you to

indicate what help, if any, you would volunteer to give the person you have been thinking about. Please answer as honestly as possible. If at all possible, please make your decisions in our study as you would in real life.

Scenario
You hear that he/she was just evicted from his/her accommodation and he/she does not have a place to stay. What (if anything) would you be willing to do in this situation? (Please check all that apply)

Nothing	□
Suggest to him/her a source of information for accommodation	□
Help him/her find a new place to live by driving him/her around for a few hours	□
Offer to have him/her stay with you for a couple of days (provided you have space)	□
Offer to have him/her come stay with you for a week (provided you have space)	□
Offer to have him/her come stay with you until he/she found a new place (provided you have space)	□
Offer to let him/her come live with you rent-free (provided you have space)	□

Continue [*Button*]

****

Participants then faced a set of adjectives describing possible emotional reactions they could have felt if they were to experience that scenario. The instructions and the first of three subsets of adjectives were presented as follows:

****

Please keep the described scenario in mind and your attention focused on [*Person X – Output field*].
Note that in order to help you in focusing on the right person we have provided you with the "*Group layout*" below, which describes your positions in the block of Lab workstations.

Group Layout

Person X	
Person Y **You**	

Imagining that you were really experiencing that scenario, please select the number that best indicates the degree to which you would have experienced each of these emotional reactions. Do not worry if you would not have felt many of these emotions; only a few may apply to a particular scenario. Be sure to select a response for each of the twenty-nine items that you will see on the next three screens.

In giving your responses please interpret the numbers as follows:

Not at all Moderately Very much so

1 2 3 4 5 6 7

	1	2	3	4	5	6	7
1. Alarmed	o	o	o	o	o	o	o
2. Grieved	o	o	o	o	o	o	o
3. Sympathetic	o	o	o	o	o	o	o
4. Intent	o	o	o	o	o	o	o
5. Soft-hearted	o	o	o	o	o	o	o
6. Troubled	o	o	o	o	o	o	o
7. Warm	o	o	o	o	o	o	o
8. Concerned	o	o	o	o	o	o	o
9. Distressed	o	o	o	o	o	o	o
10. Low-spirited	o	o	o	o	o	o	o

Confirm [*Button*]

****

The second subset of adjectives was then introduced. The exact wording is reported below:

****

Please continue by keeping the described scenario in mind and your attention focused on [*Person X – Output field*].
Note that in order to help you in focusing on the right person we have provided you with the "*Group layout*" below, which describes your positions in the block of Lab workstations.

Group Layout

Person X	
Person Y **You**	

In giving your responses please interpret the numbers as follows:

Not at all Moderately Very much so
1 2 3 4 5 6 7

	1	2	3	4	5	6	7
11. Intrigued	o	o	o	o	o	o	o
12. Compassionate	o	o	o	o	o	o	o
13. Upset	o	o	o	o	o	o	o
14. Disturbed	o	o	o	o	o	o	o
15. Tender	o	o	o	o	o	o	o
16. Worried	o	o	o	o	o	o	o
17. Moved	o	o	o	o	o	o	o
18. Disconcerted	o	o	o	o	o	o	o
19. Feeling low	o	o	o	o	o	o	o

	1	2	3	4	5	6	7
20. Perturbed	o	o	o	o	o	o	o

Confirm [*Button*]

****

The final subset of adjectives was then presented; the text read as follows:

****

Please continue by keeping the described scenario in mind and your attention focused on [*Person X – Output field*].
Note that in order to help you in focusing on the right person we have provided you with the "*Group layout*" below, which describes your positions in the block of Lab workstations.

Group Layout

Person X	
Person Y **You**	

In giving your responses please interpret the numbers as follows:

Not at all Moderately Very much so

1 2 3 4 5 6 7

	1	2	3	4	5	6	7
21. Heavy-hearted	o	o	o	o	o	o	o
22. Sorrowful	o	o	o	o	o	o	o
23. Bothered	o	o	o	o	o	o	o
24. Kind	o	o	o	o	o	o	o
25. Sad	o	o	o	o	o	o	o
26. Touched	o	o	o	o	o	o	o

	1	2	3	4	5	6	7
27. Fortunate	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
28. Guilty	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
29. Advantaged	o	o	o	o	o	o	o

Confirm [Button]

****

Tasks 3 and 4 were repetitions of Task 2 but with the focus person being one of the remaining two experimental group members. Then, Task 5 was introduced with the following text:

****

Task 5

In this task you will be asked to answer one question with regard to **your whole group**.

Continue [*Button*]

****

After the above introductory text, a We scale measurement for the whole experimental group was elicited by using the following wording:

****

Please select the appropriate number below to indicate to what extent, before this experiment, you would have used the term "WE" to characterize your group as whole.

	o	o	o	o	o	o	o	
Not at all	1	2	3	4	5	6	7	Very much so

Confirm [Button]

****

Then, Task 6 was presented to the subjects; the task introduction read as follows:

****

Task 6

In this task you will be asked to make several decisions which will affect your earnings from the experiment.

Continue [Button]

****

And Task 6 instructions were given to participants with the following text:

****

Instructions for Task 6

Task 6 is about choosing between a **lottery** and a **sure payoff**.

In the following you will go through 15 choice tasks. The lottery is the same but the sure outcome varies.

In the lottery there is a 50% chance that you win 400 points and a 50% chance that you win 0 points (the outcome of the lottery is determined by a computerized random mechanism).

On the following screen we will show you the 15 choice tasks. Please decide for each of these whether you wish to play the lottery or take the sure outcome.

Once you have made your decisions for all 15 tasks, the computer will randomly pick one of the tasks.

For this task you will either play the lottery or you will receive the sure outcome, depending on what you have chosen. The exchange rate is 1 point = 1 Penny.

Continue [*Button*]

****

Following Task 6 instructions, subjects faced the Holt-and-Laury risk elicitation task presented as follows:

****

Please decide now for any choice task whether you chose the lottery or the sure outcome!

Choice task	Lottery	Sure outcome	Your Choice	
1	50% Chance 400 Points and 50% Chance 0 Points	25 Points	Lottery	Sure outcome
2	50% Chance 400 Points and 50% Chance 0 Points	50 Points	Lottery	Sure outcome
3	50% Chance 400 Points and 50% Chance 0 Points	75 Points	Lottery ○ ○	Sure outcome
4	50% Chance 400 Points and 50% Chance 0 Points	100 Points	Lottery ○ ○	Sure outcome
5	50% Chance 400 Points and 50% Chance 0 Points	125 Points	Lottery ○ ○	Sure outcome
6	50% Chance 400 Points and 50% Chance 0 Points	150 Points	Lottery ○ ○	Sure outcome
7	50% Chance 400 Points and 50% Chance 0 Points	175 Points	Lottery ○ ○	Sure outcome
8	50% Chance 400 Points and 50% Chance 0 Points	200 Points	Lottery ○ ○	Sure outcome
9	50% Chance 400 Points and 50% Chance 0 Points	225 Points	Lottery ○ ○	Sure outcome

10	50% Chance 400 Points and 50% Chance 0 Points	250 Points	Lottery	o	o	Sure outcome
11	50% Chance 400 Points and 50% Chance 0 Points	275 Points	Lottery	o	o	Sure outcome
12	50% Chance 400 Points and 50% Chance 0 Points	300 Points	Lottery	o	o	Sure outcome
13	50% Chance 400 Points and 50% Chance 0 Points	325 Points	Lottery	o	o	Sure outcome
14	50% Chance 400 Points and 50% Chance 0 Points	350 Points	Lottery	o	o	Sure outcome
15	50% Chance 400 Points and 50% Chance 0 Points	375 Points	Lottery	o	o	Sure outcome

Please press the OK-button when you have made your decisions on all choice tasks!
OK [*Button*]

****

Figure A4 below reproduces the screenshot eliciting participants' risk attitudes.

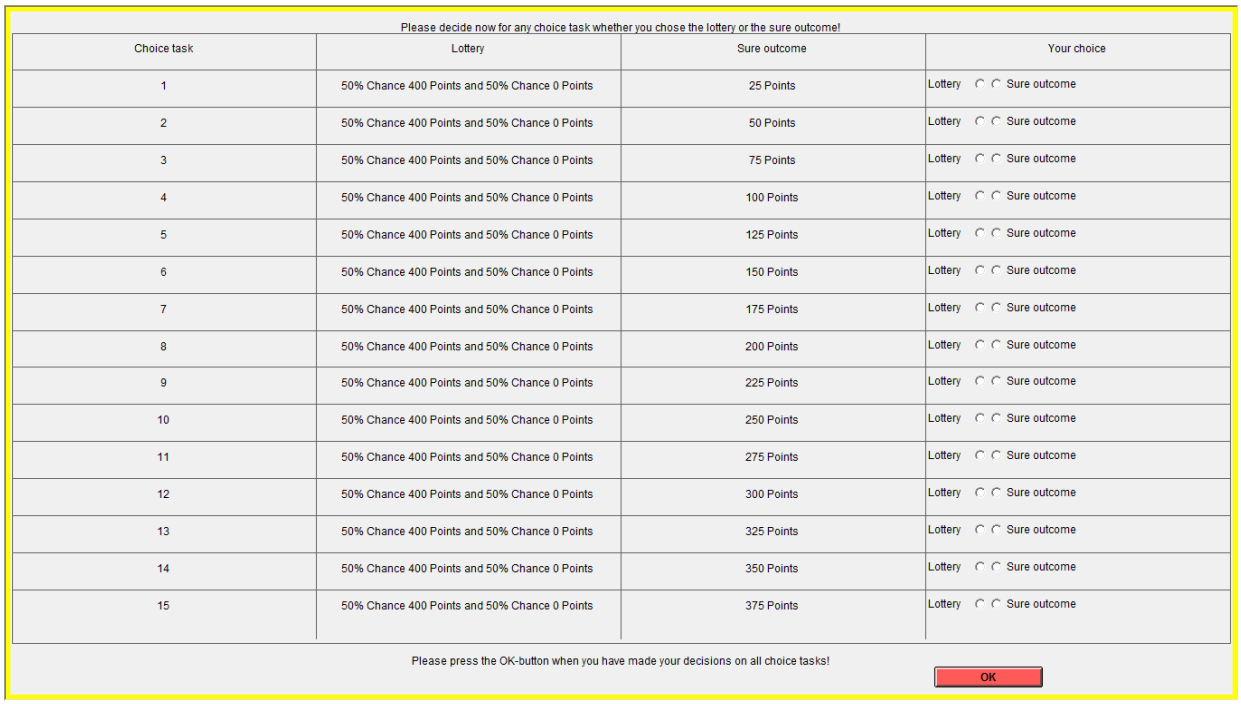


FIGURE A4. HOLT AND LAURY RISK ELICITATION METHOD, INPUT SCREEN

After participants made their choices, their computer screen would have reported their feedback as described below:

****

The computer randomly picked situation number [*Output field*] for you.
Be reminded that in the lottery there is always a 50% chance of winning 400 points and a 50% chance of winning 0 points.

Sure outcome in this situation: [*Output field*]
Your choice in this situation: [*Output field*]
Your income: [*Output field*]

Continue [*Button*]

****

Then, Task 7 was introduced with the following text:

****

Task 7

In this task you will be asked to make several decisions which will affect your earnings from the experiment.
Continue [*Button*]

****

And Task 7 instructions were presented to subjects as follows:

****

Instructions for Task 7

On the next screen you will find a list of **ten different lotteries.** In these lotteries you can earn money but you can also lose money. You have to decide for each of the ten lotteries whether you would like to participate in this lottery or not.
After you have made a decision for each of the ten lotteries whether you will participate or not, **one lottery** will be randomly selected by the computer.
In case you have chosen to **participate in the selected lottery**, the lottery will be played by the computer's random number generator, and your income will be the outcome of the lottery.
In case you have chosen **not to participate in the selected lottery**, your income will be zero.
Please click the button to go to the decision screen.
Continue to the decision screen [*Button*]

****

Following Task 7 instructions, subjects faced the loss-aversion elicitation screen, presented as follows:

****

Lottery 1	With 50% probability **you will lose 1 Pound**, and with 50% probability **you will win 6 Pounds**.	○ YES, I would like to participate in this lottery ○ NO, I do not want to participate in this lottery

Lottery 2	With 50% probability **you will lose 2 Pound**, and with 50% probability **you will win 6 Pounds**.	o YES, I would like to participate in this lottery o NO, I do not want to participate in this lottery
Lottery 3	With 50% probability **you will lose 3 Pound**, and with 50% probability **you will win 6 Pounds**.	o YES, I would like to participate in this lottery o NO, I do not want to participate in this lottery
Lottery 4	With 50% probability **you will lose 4 Pound**, and with 50% probability **you will win 6 Pounds**.	o YES, I would like to participate in this lottery o NO, I do not want to participate in this lottery
Lottery 5	With 50% probability **you will lose 5 Pound**, and with 50% probability **you will win 6 Pounds**.	o YES, I would like to participate in this lottery o NO, I do not want to participate in this lottery
Lottery 6	With 50% probability **you will lose 6 Pound**, and with 50% probability **you will win 6 Pounds**.	o YES, I would like to participate in this lottery o NO, I do not want to participate in this lottery
Lottery 7	With 50% probability **you will lose 7 Pound**, and with 50% probability **you will win 6 Pounds**.	o YES, I would like to participate in this lottery o NO, I do not want to participate in this lottery
Lottery 8	With 50% probability **you will lose 8 Pound**, and with 50% probability **you will win 6 Pounds**.	o YES, I would like to participate in this lottery o NO, I do not want to participate in this lottery
Lottery 9	With 50% probability **you will lose 9 Pound**, and with 50% probability **you will win 6 Pounds**.	o YES, I would like to participate in this lottery o NO, I do not want to participate in this lottery
Lottery 10	With 50% probability **you will lose 10 Pound**, and with 50% probability **you will win 6 Pounds**.	o YES, I would like to participate in this lottery o NO, I do not want to participate in this lottery

Please press the OK button after you have decided for each of the ten lotteries whether you would like to participate in it or not.
OK [*Button*]

****

Figure A5 below reproduces the screenshot eliciting participants' loss attitudes.

FIGURE A5. LOSS-ATTITUDE ELICITATION METHOD, INPUT SCREEN

Following their response, their computer screen provided feedback as shown below:

****

Result of Task 7
[*Output field*] has been randomly selected
You have decided [*Output field*] to participate.
You [*Output field*] the lottery.
Your income in £ is: [*Output field*]
Continue [*Button*]

****

Following Task 7 feedback, they were prompted to enter their details for payment and then to start the control tasks reported in the next subsection.

A.IV. The Control Tasks

The control tasks took the form of a post-experimental, computerized, survey. This is described in this section together with various summary statistics based on responses gathered from participants in the main experiment. (The text in squared brackets was not shown to experimental participants).

> Please answer each of the following questions as accurately as possible. Naturally your responses will be completely confidential. Your answers will be of immense value for our scientific investigation. Thank you in advance for your cooperation.

[*1.*] Your gender?
[*Male 51.92%, Female 48.08%*]

[*2.*] How old are you?
[*Mean 21.01, Std. Dev. 2.99, Median 20, Max 44, Min. 17*]

[*3.*] Nationality
[*British 46.54%, EU 10.38%, Other 43.08%*]

[*4.*] How many siblings do you have?
[*Mean 1.51, Std. Dev. 1.26, Median 1, Max 8, Min. 0*]

[*5.*] If you are a student, what is your subject?
[*Arts and Education 17.69%; Business economics 7.69%; Economics 6.54%; Engineering 21.15%; Law 3.08%; Medicine and Health Sciences 11.92%; Other Social sciences 10.00%; Politics and International Relations 4.62%; Science 17.31%*]

[*6.*] When you were 16 years of age, what was the income of your parents in comparison to other families in your country?
[*Far below average 1.92%; Below average 13.46%; Average 36.54%; Above average 41.54%; Far above average 6.54%*]

[*7.*] How large was the community where you have lived the most time of your life?
[*Up to 2'000 inhabitants 17.31%; 2'000 to 10'000 inhabitants 23.08%; 10'000 to 100'000 inhabitants 28.46%; More than 100'000 inhabitants 31.15%*]

[*8.*] How many people live in your household (please include yourself)?
[*Mean 4.34, Std. Dev. 1.37, Median 4, Max 10, Min. 1*]

[*9.*] How large is your monthly budget (expenses for accommodation already detracted)?
[*Mean 293.48, Std. Dev. 418.93, Median 200, Max 6000, Min. 0*]

[*10.*] What share of your monthly expenses you finance yourself?
[*Mean 43.23, Std. Dev. 38.45, Median 30, Max 100, Min. 0*]

[*11.*] Are you active in one of the following organizations? If so, please indicate whether you are just a member, an active member or in the board.

Sport clubs
[*No membership 32.69%, Member 31.54%, Active member 21.15%, On the board 14.62%*]

Music group
[*No membership 52.69%, Member 20.00%, Active member 10.77%, On the board 16.54%*]

Political party
[*No membership 65.77%, Member 14.23%, Active member 1.15%, On the board 18.85%*]

Lobby group (e.g. student association)
[*No membership 56.54%, Member 21.54%, Active member 6.92%, On the board 15.00%*]

Non-profit institution
[*No membership 49.62%, Member 24.62%, Active member 10.38%, On the board 15.38%*]

Other kind of clubs
[*No membership 30.77%, Member 38.46%, Active member 17.31%, On the board 13.46%*]

[*12.*] Are you religious?
[*Likert scale from 1 "Not at all religious" to 7 "Very religious": Mean 3.34, Std. Dev. 2.07, Median 3, Max. 7, Min. 1*]

[*13.*] Please indicate your political attitude in the following scale.
[*Likert scale from 1 "Left" to 7 "Right": Mean 3.81, Std. Dev. 1.38, Median 4, Max. 7, Min. 1*]

[*14.*] All things considered, how satisfied are you with your life as a whole in these days?
[*Likert scale from 1 "Not at all satisfied" to 10 "Absolutely satisfied": Mean 7.27, Std. Dev. 1.90, Median 8, Max. 10, Min. 1*]

[*15.*] How satisfied you expect to be in five years time?
[*Likert scale from 1 "Not at all satisfied" to 10 "Absolutely satisfied": Mean 8.07, Std. Dev. 1.56, Median 8, Max. 10, Min. 1*]

[*16.*] How do you see yourself? Are you generally a person who is fully willing to take risks or do you try to avoid taking risks?
[*Likert scale from 0 "Completely unwilling to take risks" to 10 "Completely willing to take risks": Mean 5.23, Std. Dev. 2.28, Median 5, Max. 10, Min. 0*]

People can behave differently in different situations. How would you rate your willingness to take risks in the following contexts?

[*17.*] How would you rate your willingness to take risks while driving a car?
[*Likert scale from 0 "Completely unwilling to take risks" to 10 "Completely*

willing to take risks": Mean 3.10, Std. Dev. 2.81, Median 2, Max. 10, Min. 0]

[*18.*] How would you rate your willingness to take risks in financial matters?
[*Likert scale from 0 "Completely unwilling to take risks" to 10 "Completely willing to take risks": Mean 4.39, Std. Dev. 2.40, Median 4, Max. 10, Min. 0*]

[*19.*] How would you rate your willingness to take risks during sports and leisure?
[*Likert scale from 0 "Completely unwilling to take risks" to 10 "Completely willing to take risks": Mean 6.22, Std. Dev. 2.68, Median 7, Max. 10, Min. 0*]

[*20.*] How would you rate your willingness to take risks in job matters?
[*Likert scale from 0 "Completely unwilling to take risks" to 10 "Completely willing to take risks": Mean 5.48, Std. Dev. 2.47, Median 6, Max. 10, Min. 0*]

[*21.*] How would you rate your willingness to take risks in health matters?
[*Likert scale from 0 "Completely unwilling to take risks" to 10 "Completely willing to take risks": Mean 3.03, Std. Dev. 2.71, Median 2, Max. 10, Min. 0*]

[*22.*] Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?
[*"Most people can be trusted" 54.23%, "Can't be too careful" 45.77*]

[*23.*] Do you think most people would try to take advantage of you if they got a chance, or would they try to be fair?
[*"Would take advantage of you" 41.54%, "Would try to be fair" 58.46%*]
[*24.*] Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?
[*"Try to be helpful" 51.92%, "Just look out for themselves" 48.08%*]

[*25.*] You can't count on strangers anymore.
[*"More or less agree" 60%, "More or less disagree" 40%*]

[*26.*] Which of the following best describes the relationship(s) between you and your partner(s) in the experiment (check all that apply)?
[*In what follows, numbers indicates the percentage of participants checking each statement: "We have no relationship at all (e.g. we just met to sign up for the experiment)"; "We are friends (e.g. we see each other under various environments and we know each other's friends)."; "We are acquaintances (e.g. we know each other but we normally don't interact)."; "I am involved in a romantic relationship with one of my partners (e.g. we are currently dating, boyfriend/girlfriend, or husband/wife)."; "We are colleagues (e.g. we see each other only at the university/work and we have little contact besides then)."; "We are family (e.g. we are siblings, cousins, etc.)"*]

[*27.*] During the last 6 months, how would you characterize the average frequency of contact between you and your partners?

[*Likert scale from 1 "No Contact" to 7 "Very Frequent Contact": Mean 5.32, Std. Dev. 2.01, Median 6, Max. 7, Min. 1*]

[*28.*] How many other participants (not including your friends) of the experiment do you know by name?
[*Mean 0.64, Std. Dev. 1.34, Median 0, Max. 10, Min. 0*]

[*29.*] Are you going to share your earnings from the experiment with your partner(s) shortly?
[*Likert scale from 1 "Certainly not" to 7 "Certainly": Mean 3.45, Std. Dev. 2.05, Median 3, Max. 7, Min. 1*]

[*30.*] Did the possibility that you can share your earnings with your partner(s) after the experiment play a role in taking your decision?
[*Likert scale from 1 "No role at all" to 7 "Decisive role": Mean 2.82, Std. Dev. 2.01, Median 2, Max. 7, Min. 1*]

[*31.*] You can rely on my data
[*Likert scale from 1 "You can not rely on my data" to 9 "You can rely on my data": Mean 7.84, Std. Dev. 1.61, Median 8, Max. 9, Min. 1*]

Your show-up fee for coming in time is 2.00 (in pounds).
Whereas, your experimental earnings are [*Output field*] (in pounds).
Hence, your total earnings in the experiment is [*Output field*] (in pounds).

Thank you very much for your participation!

Appendix B: Individual and group characteristics

B.I. Analysis of Variance Within and Between Groups

Using only F-matching data from the main experiment, Table B1 below shows the set of ANOVA and Kruskal-Wallis tests comparing the *between* and *within* group variance for the 25 characteristics elicited via the 'control tasks'.

TABLE B1—TESTING FOR EQUALITY OF VARIANCE WITHIN AND BETWEEN GROUPS (F-MATCHING DATA)

		ANOVA		K.-W. test			ANOVA		K.-W. test
#	Individual-level Variable	F	R^2	χ^2 (with ties)	#	Individual-level Variable	F	R^2	χ^2 (with ties)
1	Nationality^i	26.700***	0.897	167.807***	14	City size^i	1.750***	0.364	68.677**
2	Age^i	8.300***	0.730	155.573***	15	Trust index^i	1.640**	0.348	66.685**
3	Field of studies^i	4.780***	0.609	114.082***	16	Risk attitude^i	1.730***	0.361	64.846**
4	Gender^i	4.610***	0.601	112.323***	17	Self-financed^i	1.550**	0.336	63.338**
5	Team perception index^i	2.900***	0.486	92.868***	18	Political attitude	1.360*	0.308	60.991*
6	Religiousness^i	2.630***	0.462	82.772***	19	Empathy index^i	1.450**	0.321	60.298*
7	No. of club memberships^i	2.350***	0.434	80.388***	20	Income rank^i	1.280	0.295	60.255*
8	Prior group interactions^i	3.870***	0.558	77.187***	21	Behavioral risk attitude^i	1.580**	0.376	59.573*
9	Share intention^i	2.230***	0.421	76.650***	22	Trust strangers^i	1.140	0.271	50.704
10	Share influence^i	1.960***	0.390	76.527***	23	Future happiness^i	1.020	0.249	49.863
11	Monthly budget^i	1.056**	0.337	74.820***	24	Current happiness^i	1.060	0.257	47.546
12	No. of siblings^i	2.020***	0.397	71.930***	25	Loss attitude^i	1.000	0.260	44.501
13	Cohabitee^i	2.190***	0.417	69.942**					

Notes: Individual-level variables are ordered according to scores in the Kruskal-Wallis test (denoted "K-W" in the table). Variable specification is given below in section B.II. ***, ** and * denote significance at the 1, 5 and 10 percent levels respectively.

B.II. Individual-level Variable Construction

In this section, we explain the construction of the individual-level variables (all identified with the superscript *i*) based on the control task data. We list them in their order of appearance in Table B1. (The text in squared brackets reports the reference number of the relevant question in section A.IV).

1. *Nationality*^{i} is a categorical variable assuming values 1 (British), 2 (EU), and 3 (other). [*3.*]
2. *Age*^{i} is an interval variable measured in years. [*2.*]
3. *Field of studies*^{i} is a categorical variable taking values 0 (arts and education), 1 (economics and business), 2 (social sciences), 3 (medicine), and 4 (engineering and natural sciences). [*5.*]
4. *Gender*^{i} is a dummy variable equal to 1 for "female" and 0 for "male". [*1.*]

5. *Team perception index*[i] measures the extent to which a person would use the term "we" for the experimental group as whole. It ranges from 1 (not at all) to 7 (very much so).

6. *Religiousness*[i] indicates individual religiosity and ranges from 1 (not at all religious) to 7 (very religious). [*12.*]

7. *No. of club memberships*[i] counts the number of memberships in six possible categories of voluntary associations. It takes values from 1 (no membership at all) to 7 (membership in all six categories). [*11.*]

8. *Prior group interactions*[i] indicates the average frequency of contact with other group members in the last six months. It ranges from 1 (no contact) to 7 (very frequent contact). [*27.*]

9. *Share intention*[i] indicates the intention to share the earnings with other group members after the experiment. It ranges from 1 (certainly not) to 7 (certainly). [*29.*]

10. *Share influence*[i] documents how the possibility of sharing experimental earnings afterwards affected decisions. It ranges from 1 (certainly not) to 7 (certainly). *Share intention*[i] and *Share influence*[i] were elicited by including two questions, which follow Reuben and van Winden, (2008). We are grateful to Frans van Winden for suggesting these questions to us. [*30.*]

11. *Monthly budget*[i] reports (in pounds) the budget available per month (expenses for accommodation already detracted). [*9.*]

12. *No. of siblings*[i] is an interval variable documenting the number of siblings. [*4.*]

13. *Cohabitee*[i] is an interval variable indicating the number of people in the household (respondent included). [*8.*]

14. *City size*[i] is a categorical variable for the size of the city in which respondents spent most of their life, varying from 1 (up to 2'000 inhabitants) to 4 (more than 100'000 inhabitants). [*7.*]

15. *Trust index*[i] reports the average score of three items from the "General Social Survey" following methods from Glaeser et al. (2000). Its admissible range is from 0 to 1. [*22; 23; 24.*]

16. *Risk attitude*[i] is generated by a multi-item questionnaire (Dohmen et al., 2011) and indicates the average of six behavioral risk categories, which related to both general risk attitude as well as specific domains as for instance health matters. Its admissible range goes from 0 to 10.

17. *Self-financed*[i] indicates the percentage of the monthly expenses that is self-financed. [*10.*]

18. *Political attitude*[i] describes political alignment by ranging from 1 "left-wing" to 7 "right-wing". [*13.*]

19. *Empathy index*^{i} is generated by averaging three items (i.e., "sympathetic", "softhearted" and "compassionate") and measures the individual attitude to empathize on a scale going from 1 (not at all) to 7 (very much so).

20. *Income rank*^{i} is a categorical variable for the relative family income of the respondents, ranging from 1 (far below average) to 5 (far above average). [*6.*]

21. *Behavioral risk attitude*^{i} is generated by a standard lottery task based on Holt and Laury (2002) and measures individual propensity to risk. It assumes values from 0 (risk aversion) to 15 (risk seeking). In the statistical analyses comparing within and between group variance (e.g., see Table B1), for the variable *Behavioral risk attitude*^{i} only 236 observations were used rather than 260, because 24 participants displayed choice inconsistency. (This lottery task can be found in Task 6, section A.III).

22. *Trust strangers*^{i} is a dummy variable equal to 1 when respondents state they count on strangers and 0 otherwise. [*25.*]

23. *Future happiness*^{i} is a categorical variable documenting the life satisfaction expected in five years time. It varies from 1 (not at all satisfied) to 10 (absolutely satisfied). [*15.*]

24. *Current happiness*^{i} is a categorical variable indicating current life satisfaction. It takes values between 1 (not at all satisfied) and 10 (absolutely satisfied). [*14.*]

25. *Loss attitude*^{i} was elicited via a task based on Gächter, Johnson and Herrmann (2010). It measures individual propensity to loss. Its admissible range is from 0 (loss aversion) to 10 (loss seeking). In the statistical analysis comparing within and between group variance (e.g., see Table B1), for the variable *Loss attitude*^{i} only 245 observations were used rather than 260, because 15 participants displayed choice inconsistency. (This lottery task can be found in Task 7, section A.III).

B.III. Group-level Variable Construction

The group-level variables have labels corresponding with the relevant individual-level variables from which we construct them, but the superscript *i* has been dropped. The group-level variables are constructed as follows. *Nationality* reports the highest proportion of group members in one of three elicited categories (i.e., British, EU, others) ranging from 0.5 (heterogeneous group) to 1 (homogenous group). *Age* is the age range of each group (e.g., if the youngest group member is 19 years old while the oldest is 23, *Age* will be equal to 4). *Gender* captures group homogeneity ranging from 0.5

(heterogeneous) to 1 (homogenous). All remaining variables are group means of the corresponding individual-level variables.

Appendix C: Group Cohesion and Background Characteristics

Table C1 below shows the Spearman correlation coefficients between group cohesion and each of the control variables featuring in the econometric models reported in Table 2 of the paper.

TABLE C1—CORRELATION COEFFICIENTS BETWEEN GROUP COHESION AND CONTROL VARIABLES

	Group Cohesion		Group Cohesion		Group Cohesion
Gender	0.453***	Monthly budget	-0.201	Share intention	0.290**
Age	-0.623***	Self-financed	-0.013	Share influence	0.316***
Field of studies	0.307**	No. of club memberships	-0.214*	Empathy index	0.560***
Nationality	0.689***	Prior group interactions	0.743***	Team perception	0.864***
No. of siblings	0.141	Religiousness	-0.269**	Trust index	0.017
City size	-0.266**	Political attitude	-0.304**	Risk attitude	0.150
Cohabitee	0.138				

Notes: Correlation coefficients are calculated by using group-level data from Period 1. Variable definition and construction are in section B.III. ***, ** and * denote significance at the 1, 5 and 10 percent levels respectively.

Appendix D: Two-phase Experiment: Additional Graphical Analysis

As reported in our paper, we discovered an intriguing difference between the two-phase experiment data and the main experiment, which emerges when we partition the data according to levels of group cohesion.

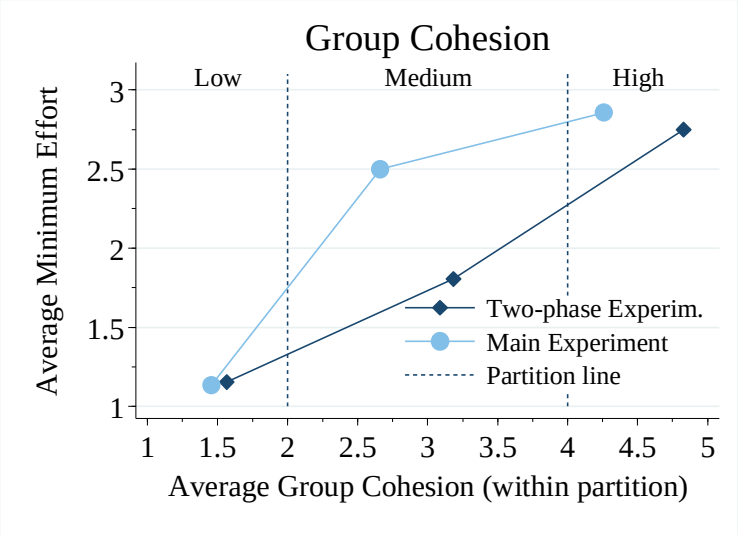


FIGURE D1. MINIMUM EFFORT AND GROUP COHESION IN TWO-PHASE EXPERIMENT AND MAIN EXPERIMENT (GROUP-LEVEL DATA FROM PERIOD 8)

Note: For comparability, robustness check is based on week 2 group cohesion. "Low Cohesion": *group cohesion* $\in$ [1, 2]; "Medium Cohesion": *group cohesion* $\in$ (2, 4]; "High Cohesion": *group cohesion* $\in$ (4, 7].

Figure D1 constructs two piecewise-linear approximations to capture the relationship between group cohesion and minimum effort, separately for the main experiment and the two-phase experiment. In this analysis, the two-phase experiment group cohesion is the average of our group cohesion statistic in week 2 but averaged across groups within partition. Minimum effort data are from period 8 of the weak-link game. In each case, the piece-wise linear approximations are composed by partitioning the data into subsets for *low*, *medium* and *high* group cohesion (for details on the partitions see note to Fig. D1). We then plot mean group cohesion against the mean of minimum effort for each partition of each data set. The figure provides graphical illustration of the difference between the two-phase and the main experiment for medium cohesion groups. We discuss the interpretation of this difference in the main paper.

REFERENCES

Dohmen, Thomas, Falk, Armin, Huffman, David, Sunde, Uwe, Schupp, Jürgen, and Wagner, Gert G. 2011. "Individual Risk Attitudes: Measurement, Determinants, and Behavioral Consequences." *Journal of the European Economic Association*, 9(3): 522-50.

Gächter, Simon, Johnson, Eric, and Herrmann, Andreas. 2010. "Individual-Level Loss Aversion in Riskless and Risky Choices." CeDEx Discussion Paper No. 2010-20, University of Nottingham.

Glaeser, Edward L., Laibson, David I., Scheinkman, Jose A., and Soutter, Christine L. 2000. "Measuring Trust." *Quarterly Journal of Economics*, 115(3): 811-46.

Holt, Charles A., and Laury, Susan K. 2002. "Risk Aversion and Incentive Effects." *American Economic Review*, 92(5): 1644-55.

Reuben, Ernesto, and van Winden, Frans. 2008. "Social Ties and Coordination on Negative Reciprocity: The Role of Affect." *Journal of Public Economics*, 92(1-2): 34-53.