CENTRE FOR DECISION RESEARCH & EXPERIMENTAL ECONOMICS

University of Nottingham
UK | CHINA | MALAYSIA

Discussion Paper No. 2019-16

Simon Gächter, Chris Starmer and Fabio Tufano

December 2019

The Surprising Capacity of the Company You Keep: Revealing Group Cohesion as a Powerful Factor of Team Production

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

The Surprising Capacity of the Company You Keep: Revealing Group Cohesion as a Powerful Factor of Team Production

Simon Gächter, Chris Starmer and Fabio Tufano*

*University of Nottingham***

December 20, 2019

We introduce the concept of "group cohesion" to study the economic consequences of social relationships in team production. We measure group cohesion, adapting the "oneness scale" from psychology to group level. A series of experiments, including a pre-registered replication, reveals that higher cohesion groups are more likely to achieve Pareto-superior outcomes in weak-link coordination games. Judged against benchmarks, the effects of cohesion are economically large. We identify beliefs rather than social preferences as a primary mechanism explaining the effects of cohesion. Our comprehensive evidence establishes group cohesion as a powerful production factor and a useful new tool of economic research.

JEL Classification: C92, D03

* Corresponding author (fabio.tufano@nottingham.ac.uk). ** Address: School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, UK. We are grateful to Robert Cialdini and John K. Maner for sharing their experimental tools. We thank Antonio Aréchar, Michalis Drouvelis, Francesco Fallucchi, Ernesto Gavassa Perez, Jose Vicente Guinot Saporta, 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, Leonardo Bursztyn, Colin Camerer, Gary Charness, Yan Chen, David Cooper, Vincent Crawford, Robin Cubitt, Enrique Fatas, Sourafel Girma, John Hillas, John List, Muriel Niederle, Pietro Ortoleva, Carla Rampichini, Al Roth, Marie-Claire Villeval, Roberto Weber, Frans van Winden and Leat Yariv. We also thank participants at various conferences, workshops and seminars for comments and suggestions. The work was supported by the Economic and Social Research Council (grant numbers ES/K002201/1, ES/P008976/1). Tufano thanks also MIUR for financial support. Gächter received support from the European Research Council grant ERC-AdG 295707 COOPERATION. Research funding from the University of Nottingham is also gratefully acknowledged. The research reported in this paper was approved by the Research Ethics Committee of the Nottingham School of Economics. The authors declare no relevant or material financial interests that relate to the research described in this paper.

I. Introduction

A vast array of economic and social activity occurs in groups and teams. 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 the role of incentives for team performance (e.g., Gibbons and Roberts (2013)), not much is known about the productive potential of *social relationships* among group members. In what follows, we report an extensive program of conceptual and experimental research building from the arguably plausible but, so far, barely considered idea that the psychological qualities of social relationships within households, firms and other organizations, collectively constitute an economically important factor of team production.

Our goal in this paper is to develop an understanding of the hidden potential of social relationships for team performance, via the introduction and careful study of a new concept: *group cohesion*. We proceed in three steps. The first is to make the quality of social relationships in real existing groups measurable by characterizing any group in terms of its group cohesion index. The group cohesion index is a function of how socially close members of a group perceive each other to be according to a psychologically well-founded instrument, which we also tested and validated in related research (Gächter, Starmer, and Tufano (2015)). Our second step is to demonstrate the power of group cohesion in a large-scale program of experiments investigating team production. Because our research question demands real groups, we use groups of friends, as well as standard randomly assembled groups as participants. We show that group cohesion has positive, economically large, and replicable effects on group performance. Our third step is to study two potential mechanisms of how group cohesion can influence group performance: social preferences and beliefs about others' behavior. Our experiments and econometric analyses identify beliefs rather than social preferences as the primary mechanism.

Since group cohesion is a novel concept in economics, we proceed by first providing a broader rationale for the concept and its measurement. Our starting point is that members of *any* real human group inevitably have some relationships to other group members: for example, to begin with a very simple idea which we later operationalize, 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 previous research (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 average of the individual minimum oneness ratings (details in Section III). As we will show, this seemingly modest measurement innovation generates a tool with considerable predictive power.

In a series of seven experiments (summarized in the Supplemental Material, SM – Table SM1.1), we demonstrate that group cohesion is an economically important determinant of group outcomes. A key feature of our setup is that we study the behavior of *real* groups, not artificially created ones, achieved by recruiting groups of friends to participate. This approach – as we show in detail in Section III – allows us to observe real closeness ties within real groups. Real groups have sociological and psychological characteristics that are absent (by construction) in groups set up on the spot in the lab, including groups constructed using minimal group manipulations (Goette, Huffman, and Meier (2012)).[1]

Bringing real groups to the lab is a departure from classic lab experiments which might, initially, raise eyebrows among those who presume that (at least approximate) anonymity is a sine qua non principle for experimental games, required to avoid the shadow of the future 'infecting' strategic behavior in the lab. We aim to convince readers otherwise. A key rationale for our approach comes from the fact that real groups, and the real relationships that have developed within them, are our object of study. While we recognize that working with real groups does create some methodological challenges and issues of interpretation, we address

[1] In the taxonomy of Charness, Gneezy, and Kuhn (2013), our experiments would classify as "extra-laboratory experiments," while in that of Harrison and List (2004) they are "artefactual field experiments."

them directly within our studies and find reassuring results (see Section VI and VII). As such, we see it as part of our contribution to demonstrate the practicability of, and the considerable potential value in, lab-experimental research with real groups.

Our workhorse to study group outcomes is a weak-link coordination game chosen partly because it captures economically interesting problems endemic to organizations and teams (e.g., Knez and Camerer (1994); Camerer and Weber (2013)). Our version of the weak-link game is inspired by Brandts and Cooper (2006) and explained in detail in Section IV. 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 (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 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).

In Section V, we identify clear, strong and stable relationships between group cohesion and behavior in our experiments. While the likelihood of coordinating on some equilibrium appears largely independent of the level of group cohesion, it plays a major role in equilibrium selection: low cohesion groups usually descend rapidly to Pareto-inferior equilibria; high cohesion groups fare much better. High group cohesion appears necessary (though not sufficient) for achieving Pareto-superior outcomes. We confirm that these patterns replicate via a pre-registered experiment independently conducted at the University of Birmingham (Study 2, Table SM1.1). Further experiments show that our results are robust to the timing of oneness measurement (before or after play of the weak-link game).

In Section VI, regression analysis demonstrates that group cohesion is a reliable predictor of group minimum effort and gives rise to distinct dynamics in individual-level behavior. Our modelling also shows that predicted effects of cohesion are very large: low cohesion groups are almost certain to collapse to minimum effort, while maximally cohesive groups are almost certain to achieve Pareto optimality.

In Section VII, we report two further sets of experimental treatments designed to benchmark the effort-enhancing effects of group cohesion: we establish that large financial incentives are needed to match the levels of minimum effort expected for high cohesion groups. We also demonstrate that inducing social preferences among strangers (by forcing them to share their

earnings), induces behavior consistent with mid-range cohesion groups, but that uncertainty about the likelihood of sharing quickly undermines the impact of sharing rules.

In Section VIII, we turn to an explanation of our results. A rational choice perspective suggests two natural channels through which group cohesion could operate: it might affect some combination of group members' social preferences, and their beliefs about the effort levels their fellow group members will choose. Considering *social preferences*, it is plausible to assume that members of highly cohesive groups place more weight on each other's earnings.[2] In our weak-link game, if players do draw utility from each other's earnings, this reduces strategic risk and fosters coordination on Pareto-superior equilibria.[3] Because the weak-link game we study has multiple equilibria, *beliefs* about other's efforts are crucial for own effort and hence the group outcome (Knez and Camerer (1994)). 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); Kets and Sandroni (2019)). These 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, and little implicit learning to draw on from shared situations.

We probe these possibilities using mediation analysis (as used by, e.g., Kosse et al. (forthcoming)) and the data from the pre-registered replication study, which also elicited beliefs and a measure of social preferences. Our results confirm that group cohesion has a large and highly significant effect (explaining around 30% of the total effect on effort) and that the bulk of this effect operates via the channel of beliefs.

Based on our large body of results, we believe we have established proof of concept for group cohesion as both an important factor of team production and a useful new tool of economic research with considerable predictive power and potential for wide-ranging application to both theoretical and applied economics.

II. 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. As we will

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

[3] Chen and Chen (2011) develop a model in which social preferences reduce the strategic risk in coordination games. They also provide experimental evidence that artificially-induced group identities foster coordination on higher Pareto-ranked equilibria.

explain below, group cohesion builds on the concepts of 'relationship closeness' and 'oneness'. These concepts are firmly established in the psychology literature (see next section) but are barely considered in economics, with the possible exception of 'social distance'. Social distance is juxtaposed to complete anonymity and manipulated experimentally by giving participants cues about the identity of other individuals, for instance, via visual identification (Bohnet and Frey (1999)) or via their names (Charness and Gneezy (2008)); the closeness of the relationship is not explicitly measured. By contrast, we measure the closeness of relationships between group members and derive the concept of group cohesion from it. To our knowledge, this is an entirely new approach in economics.[4] 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 – following seminal papers by van Huyck, Battalio, and Beil (1990, 1991) – 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 organizational or societal culture (Weber and Camerer (2003); Engelmann and Normann (2010)). By studying a fixed weak-link game, we keep structural features constant and 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)) that includes studies investigating interactions among real-world friends (e.g., Reuben and van Winden (2008); Leider et al. (2009); Brañas-Garza et al. (2010)); the role of identity in organizations, in particular (e.g., Gibbons (2010); Akerlof and Kranton (2005); Ashraf and Bandiera (2018)); and the importance of understanding social-psychological dimensions of employment relationships, more generally (e.g., Baron and Kreps (2013)). The papers most closely related

[4] There is 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.

to ours in this strand of literature are Charness, Rigotti, and Rustichini (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. In a weak-link game context, Chen and Chen (2011) develop an identity-contingent social-preference model and show, using artificially-induced group identities, that an enhanced group identity can foster the selection of Pareto-superior equilibria. In contrast to both papers, we do not focus on the effects on game behavior of ("near") minimal group identity manipulations, which are regarded by some scholars as akin to framing manipulations (e.g., Guala and Filippin (2017)). As Goette, Huffman, and Meier (2012) forcefully argue, and as we will show empirically in Section III.B, real groups are more than just framings or labels: real groups have sociological and psychological characteristics (which are absent from minimal groups by construction) that influence group cohesion in a predictable way. Moreover, we will also show that group cohesion influences social preferences and in particular beliefs, which in our setting predict cooperation better than social preferences.

III. Group Cohesion in Real Groups

Group cohesion is a novel concept in economics and therefore we devote subsection A to explaining the nature of the concept, its roots in an established psychological literature and how we draw on this to develop a new measurement tool. In subsection B, we will provide evidence that our concept of group cohesion reflects, as relevant literature would lead us to expect, psychological and sociological properties of real groups.

A. Group Cohesion: psychological foundations and measurement

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 focus person (Aron, Aron, and Smollan (1992)). Our strategy builds on this literature and extends it 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 feel, collectively, 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 increases with people's *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" to measure in detail the frequency of interactions, diversity of jointly undertaken activities and the influence a pair exerts on each other.

While the Relationship Closeness Inventory is fine-grained, it is not practical for many purposes. To provide a handy measurement technique, in a highly cited paper, Aron, Aron, and Smollan (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, Aron, and Smollan (1992), p. 598). Essentially, it measures relationship closeness without examining its detailed determinants of frequency of interactions, diversity of activities and strength of mutual influence.

Aron, et al. provide statistical evidence that the IOS scale is successful in tracking key dimensions of broader relationship closeness: in Fig. 1a people pick the more overlapping pair of circles the more they report that they interact frequently, share diverse activities with each other, and perceive a stronger influence on one another. The IOS scale also has the decisive advantage of being intuitive and very simple to implement.

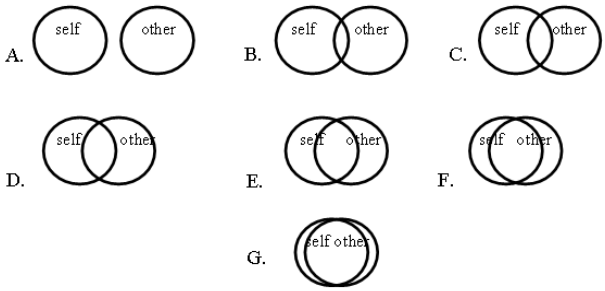
Subsequent research, most notably by Starzyk et al. (2006), developed an 18-item "Personal Acquaintance Measure" intended for application to a wider range of relationships including acquaintances. Starzyk et al. found that their measure correlates strongly with the IOS scale. Together, these results make the IOS scale a very promising tool for our purposes.

Our research question relies critically on the IOS scale. In an effort to ascertain the psychometric validity of the IOS scale for a whole range of relationships (from strangers to acquaintances, to friends and close friends), in a companion paper (Gächter, Starmer, and Tufano (2015)) we replicated Aron, et al.'s IOS scale and the following questionnaire-based measures: the Relationship Closeness Inventory (Berscheid, Snyder, and Omoto (1989)), the Personal Acquaintance Measure (Starzyk et al. (2006)) and the Loving and Liking scale (Rubin

(1970), intended to measure the evaluation and affection of a target person). We used 772 workers on Amazon Mechanical Turk as subjects.

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

FIG. 1.—Oneness elicitation as explained to the participants.

Our results replicate remarkably closely the correlations of the IOS scale that Aron, et al. report with the Relationship Closeness Inventory scales. The IOS scale correlations with the Personal Acquaintance Measure scales are also very similar to those reported by Starzyk et al. (2006). The IOS scale also varies highly significantly with the strength of the relationship (lowest for acquaintances, medium for friends, and highest for close friends) and with the Loving and Liking Scale (Rubin (1970)).

In Gächter, et al. (2015) 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 that has the properties claimed by Aron, et al.

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, Starmer, and Tufano (2015) we confirmed Cialdini et al.'s claim that oneness correlates slightly better with the questionnaire-based measures than the IOS scale alone and hence we use the oneness scale for our analysis.

In our experiments, we deployed the oneness scale as follows (wider procedural details are in Section IV). Subjects participated as groups of four and each one rated three other visually identified group members, separately and privately, by responding to the IOS and We scales as depicted in Fig. 1; group members never received feedback about each other's ratings, and they were aware of that. 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* the weakest relationship link in a group is. 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 discuss sensitivity to different ways of constructing the statistic in Section VI.B).

B. Naturally Occurring Groups and their Psychological and Sociological Properties

A key design objective was to import "real", that is, naturally occurring, groups into the lab and to characterize them with the group cohesion variable. To this end, we exploit data from experiments in which we recruited groups of four friends who were then either matched into new groups of four members (Non-friends, N-matching) or kept together as friends to proceed to the experimental tasks (Friends, F-matching); details and further rationale are in Section IV. Since group cohesion is meant to be a statistic capturing psychological dimensions of real social relationships, with the latter depending on objectively measurable features of social relationships (e.g., frequency and diversity of interactions), we should expect group cohesions to vary in predictable ways, with tangible features of our groups. So, as a preliminary test of our construct's validity, we make and test four predictions based on this presumption. Specifically, we predict that: (1) existing groups will have higher group cohesion than newly formed ones; (2) in F-matching groups, the variation in oneness between groups is bigger than the variation within groups of friends; we do not expect such a difference in N-matching groups; (3) real groups have self-selected members who are predicted to associate according

to their similarity on various socio-demographic characteristics ('homophily', see also Baccara and Yariv (2013)); (4) group cohesion increases in homophily.

To test the first prediction, Fig. 2 plots the distribution of our measure of group cohesion separately for N- and F-matching. As predicted, 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 bi-lateral social relationships.

Furthermore, Fig. 2 shows that there is a large cross-group variation in group cohesion. Group cohesion scores range from a minimum of 1 to a maximum of 5.5. Within the F-matching, the range is between 2.0 and 5.5, whereas in N-matching it is between 1.0 and 3.25. As expected, in F-matching groups, the variation of oneness ratings is larger between groups than within groups (Kruskal-Wallis test: $\chi^2(46) = 77.1$, $p < 0.003$), whereas this is not the case in N-matching groups ($\chi^2(17) = 13.4$, $p = 0.711$). This confirms prediction 2 and means that our environment provides good scope for observing effects of variation in group cohesion on group behavior, if such effects exist.

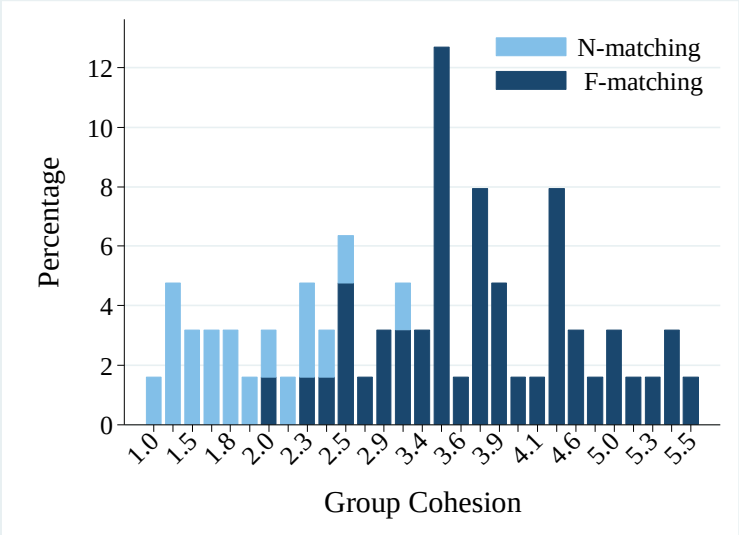


FIG. 2.—The distribution of group cohesion under F- and N-matching. The N-matching bars are stacked over the F-matching ones.

According to prediction 3 members of self-assorted real groups should be more similar to one another with regard to socio-demographic characteristics than members of different groups (Baccara and Yariv (2013)). This within-group homophily is also predicted by sociological literature on friendship ("like-befriends-like"): people tend to become friends with people who

share similar socio-demographic backgrounds, attitudes and preferences (e.g., McPherson, Smith-Lovin, and Cook (2001)). Hence, for F-matching groups, we should find greater variation of socio-demographic and attitudinal characteristics *between* groups than *within* groups. We test this using individual-level data generated from post-experimental questionnaires. Based on the non-parametric Kruskal-Wallis test, out of twenty-five variables measured in the questionnaire, 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 in Table SM2.1). By contrast, and as expected, corresponding analysis for N-matching groups reveals no significant differences for any characteristic (all $p > 0.093$), comparing within and between group homogeneity. Thus, the real groups we have recruited reflect the sociological properties predicted from relevant literature. This confirms prediction 3.

To test prediction 4 we construct an *Homophily index* that increases with the similarity of group members on each of the fifteen variables we measured to capture tangible sociological features of group members.[5] These variables are a mix of exogenous characteristics (e.g., gender, age, and nationality) but also endogenous variables such as memberships in clubs. We include those because relationship theory predicts that closeness increases with shared activities. Details on all variables are in the supplemental material (section SM3). An OLS regression of Group Cohesion on the Homophily Index produces a highly significant coefficient ($p < 0.001$) with an R^2 of 0.36. In sum, group cohesion is a very simple, intuitive and portable group-level statistic that is also a psychologically and sociologically meaningful summary of the social relationships of group members. In the remainder of this paper we study how and why group cohesion influences team performance.

IV. Experimental Setup

A. The Test Environment: The Weak-link Game

Our workhorse for studying team performance is the so called weak-link game. Since the seminal papers by van Huyck et al. (1990, 1991), 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

[5] Fifteen variables out of twenty-five variables measured in the questionnaire are used to construct the Homophily Index. Two variables were not used given that they were only elicited in Study 1. The remaining eight variables are used as controls either in Table 2 below or in the additional econometric analysis reported in the supplemental material (see SM6).

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 payoff matrix as generated by the payoff function π_i.

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). 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 two primary reasons. First, Brandts and Cooper (2006, 2007) have established that this environment is "harsh" in the sense that, in the absence of aids to cooperation (e.g., communication), behavior rapidly and reliably converges

to the lowest Pareto-ranked equilibrium.[6] Second, Brandts and Cooper (2006) 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 study described in Section VII.

B. Sampling Strategy and Sequence of Events

Our goal was to study real groups, which we achieved 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 F-matching (47 groups) or the N-matching (18 groups). In the F-matching, each quartet of acquaintances was allocated to the *same* group ("Friends"). By contrast, in the N-matching, each set of four acquaintances was split up so that each became a member of a *different* group ("Non-friends"). 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 achieved the desired variation in pre-existing cohesion across groups (Fig. 2).

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-subject experiments (172 new participants; 43 groups, 27 F-matching, 16 N-matching) over two weeks. We refer to these as our "two-week experiments" (see Table SM1.1). 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

[6] When the threshold marginal cost mc^{thr} (which is equal to the marginal return of minimal effort – mr – divided by the number of players) is lower (resp. higher) than the individual marginal cost mc, participants tend to converge on the lowest (resp. highest) Pareto-ranked equilibrium (e.g., Anderson, Goeree, and Holt (2001); Chen and Chen (2011)). It is therefore possible to use the ratio $h=mc^{thr}/mc$ as an indicator of the harshness of the game environment: the lower h is, the harsher is the environment. Consider for instance that in Chen and Chen (2011) $mc^{thr}=mr/n=1/2=0.5$ while $mc=0.75$, resulting in $h=0.67$; in our case $mc^{thr}=mr/n=(10\ b)/4= 60/4=15$ while $mc=50$, resulting in $h=0.30$. Therefore, our environment is harsher than Chen and Chen's.

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, $\beta = -0.032$, $z = -0.28$, $p = 0.783$).[7] 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 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. 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

[7] 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.

the IOS scale and the We scale (Fig. 1) for the relevant focus person. The experimental instructions are documented in the supplemental material (section SM4).

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.

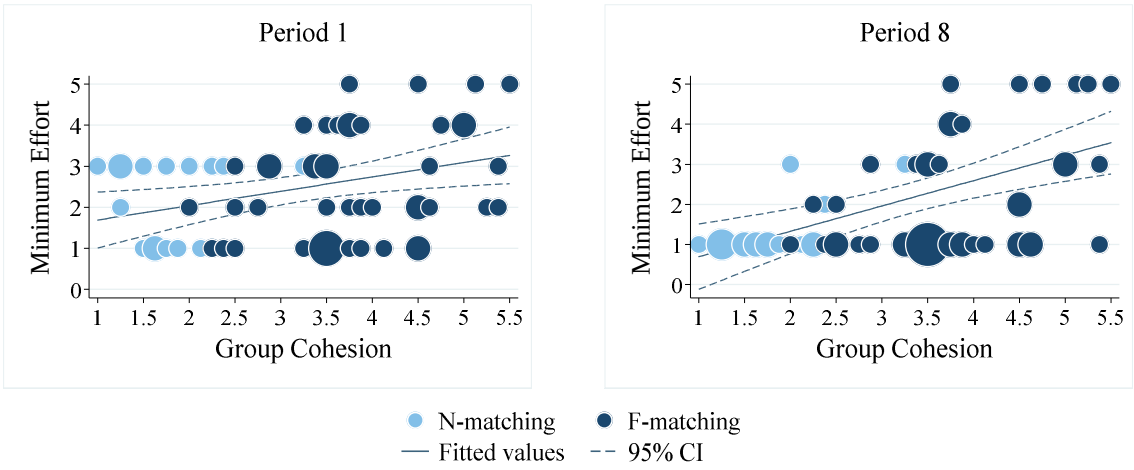
V. Establishing Group Cohesion as a Factor of Team Production

Before presenting our primary results, we note that our experimental environment is "harsh", as intended, 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 cooperation failure is robust to game length and a non-student subject pool.[8] These results confirm existing evidence by Brandts and Cooper (2006, 2007) on the harshness of our experimental environment and show that there is ample scope for improvement in cooperation, potentially in response to variation in the group cohesion measure we study here.

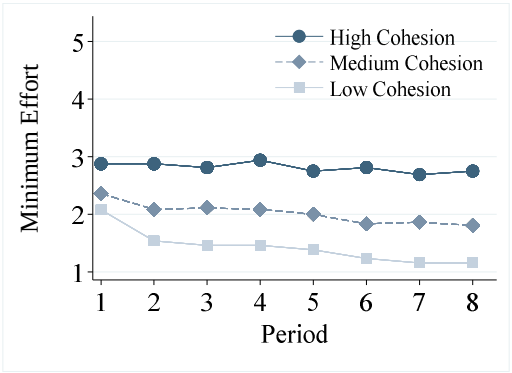
A. The Impacts of Group Cohesion on Minimum Effort and Wasted Effort

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.

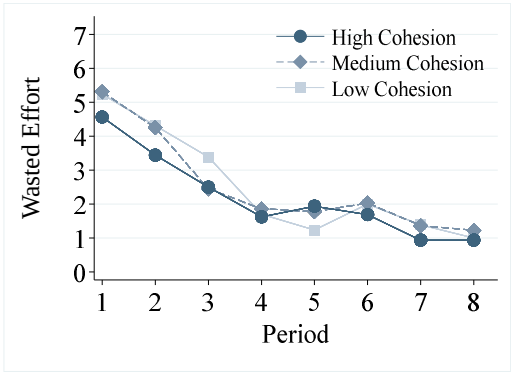
[8] To test whether a longer game might promote coordination on better equilibria, 32 fresh participants, recruited individually, played the game of Table 1 for 50 periods in fixed groups of four anonymous members (details in Table SM1.2). Six out of the eight groups were trapped into the lowest Pareto-ranked 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 further test, 44 fresh participants, recruited at the Joint Officer Training Program (JOTP) of the Swiss Army, played the game for 8 periods (details in Table SM1.2). These participants knew other group members were from the JOTP but not their exact identities. 73% of these groups collapsed to the Pareto-worst equilibrium from which we conclude that general knowledge about aspects of the identity of other group members is not enough to support cooperation.



3a. The link between group cohesion and group-minimum effort in Period 1 and 8



3b. Group minima across periods



3c. Wasted effort

FIG. 3.—Group Cohesion, Minimum Effort and the Dynamics of Coordination. Fig. 3a: 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. Fig. 3b and 3c: "Low Cohesion" Partition (13 groups): group cohesion $\in$ [1, 2]; "Medium Cohesion" Partition (36 groups): group cohesion $\in$ (2, 4]; "High Cohesion" Partition (16 groups): group cohesion $\in$ (4, 7]. Fig. 3b: average group minimum effort over time. Fig. 3c: wasted effort per period, calculated as the sum of efforts in a group above the group minimum, averaged across groups.

To further examine the dynamics suggested by Fig. 3a, we separate the full set of 65 groups into three subsets of "low", "medium" and "high" cohesiveness groups (for details of partitions see Fig. 3 caption). Fig. 3b reveals marked differences in the 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. 3c 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.[9] The analyses of Figs. 3b and 3c suggest that group cohesion primarily facilitates cooperation (fostering decisions consistent with higher ranked equilibria), with 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.

A natural question to ask is whether our results are robust to the timing of the oneness elicitation. We use the data generated by our "two-week" experiment (where oneness is also elicited one week before the weak-link experiment, see Section IV.B) to conduct a simple but informative check comparing average minimum effort across experiments (original vs two-week experiment) using the partitions for group cohesiveness (i.e., low, medium, and high) introduced in Fig. 3b. These tests show that for both low and high cohesion groups, achieved minimum effort is the same in the main and the two-week experiment; for groups with mid-range cohesion, minimum effort is somewhat higher for the two-week experiment. We conclude that our results are robust to the timing of the oneness elicitation. Additional supporting analysis can be found in the supplemental material (section SM5).

While the results documented in this subsection are encouraging, they are also novel. Therefore, replicability is a first-order methodological issue (e.g., Camerer, Dreber, and Johannesson (2019)) to be resolved in order to provide solid facts on which to rest further analyses on the economic consequences of group cohesion. We therefore replicated the experiments reported in Section IV.A and report the results in the following sub-section.

B. A Pre-Registered Replication

In the following, we sometimes use "Study 2" as a convenient label for the replication study and refer back to the original study as "Study 1". To provide a credible replication, we pre-registered the experiments[10] for Study 2 and we hired an outside contractor (the University of Birmingham Experimental Economics Laboratory (BEEL)) to run them for us. We provided the experimental protocol, software and instructions, but we were not involved in data collection. This has two advantages: there cannot be a "Nottingham experimenter effect"

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

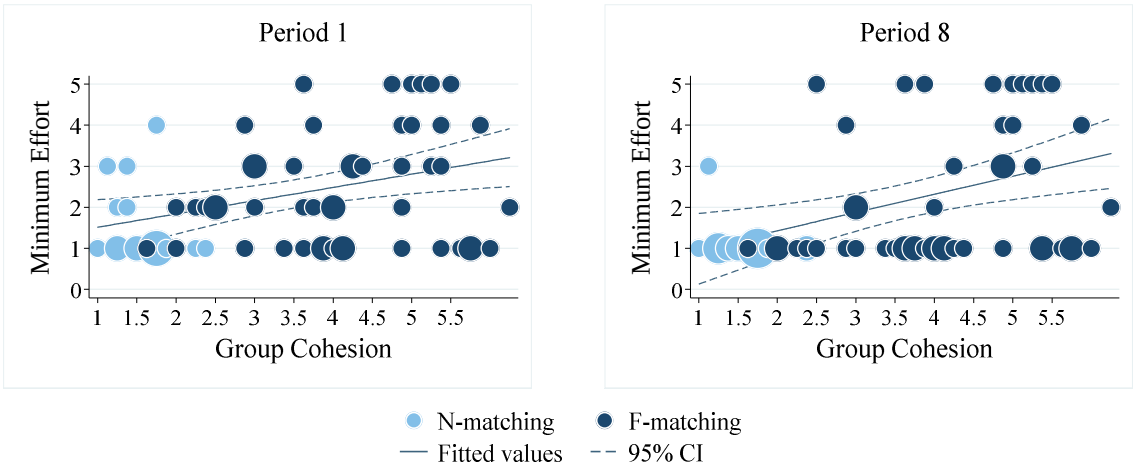
[10] Registration number AEARCTR-0003566. See https://www.socialscienceregistry.org/trials/3566. We planned to collect data from 70 independent experimental groups for a total of 280 participants. Due to no-shows, we fell short of 1 experimental group (4 participants) in the F-matching.

because the study was run independently by BEEL experimenters and the replication happened with a new subject pool from another university. We ran the exact same protocol as described above (with 276 participants; 49 F-matching groups and 20 N-matching ones) except that, with a view to probing the causes of the relationship identified in Study 1, we introduced two further sets of measurements. First, we elicited subjects' beliefs in each round of the weak-link game about the minimum effort in their group. Second, the post-experimental questionnaire also included an incentivized elicitation of "Social Value Orientation" (Murphy and Ackermann (2014)) as a simple measure of group social preferences. We discuss the details of these measures and the associated results in Section VIII.

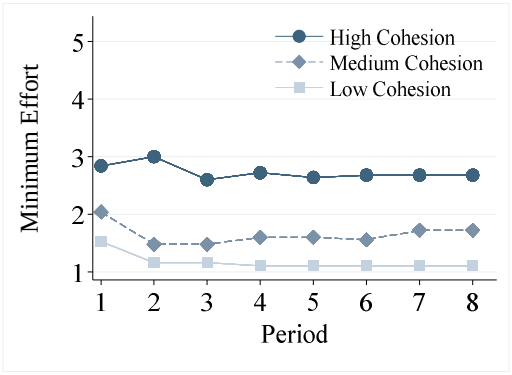
Before going to the main results of Study 2, it is worth noticing that Study 2 closely replicated the findings on (i) the distributions of group cohesion scores between the N- and F-matching; (ii) homophily, that is smaller within-group variance than between-group on the large majority of variables; and (iii) the links between group cohesion and the homophily index. Details are in Fig. SM2.1 and Table SM2.1.

Fig. 4 describes the main results of Study 2. A comparison with the corresponding Fig. 3 for Study 1 reveals that, qualitatively, the results are remarkably similar. Panel 4a replicates the positive relationships between group cohesion and minimum effort though with the difference that, in the replication, the relationship is strongly significant for both periods. Fig. 4b confirms the ability of higher cohesion groups to achieve and sustain higher minimum effort levels over time while Fig. 4c confirms the finding that cohesion levels do not substantially affect the dynamics of wasted effort.[11]

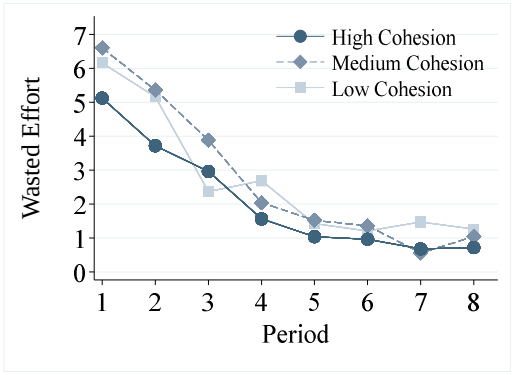
[11] As for Study 1, we find only a weakly significant relationship between (average) Study-2 group level wasted effort and group cohesion (Spearman's ρ = -0.209, p = 0.085; n = 69).



4a. The link between group cohesion and group-minimum effort in Period 1 and 8



4b. Group minima across periods



4c. Wasted effort

FIG. 4.—Study 2: pre-registered replication independently conducted at the University of Birmingham. Fig. 4a: Size of symbols proportional to the number of observations; in both Period 1 and 8, three N-matching observations are not displayed because they coincide with F-matching circles with coordinates (2, 1), (2.375, 2) and (2.875, 1); OLS Regression (69 groups), Period 1: $\beta = 0.321$ (se = 0.099, $p = 0.014$, $R^2 = 0.135$); Period 8 data: $\beta = 0.405$ (se = 0.108, $p < 0.001$, R^2=0.175); an ordered probit estimation generates qualitatively similar results. Fig. 4b and 4c: "Low Cohesion" Partition (13 groups): group cohesion $\in [1, 2]$; "Medium Cohesion" Partition (36 groups): group cohesion $\in (2, 4]$; "High Cohesion" Partition (16 groups): group cohesion $\in (4, 7]$. 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.

In sum, the results of Study 2 confirm that group cohesion has a replicable positive impact on cooperation in the weak-link game. In subsequent sections of the paper, we examine the power of this relationship and its determinants.

VI. The Predictive Power of Group Cohesion for Minimum Effort

The combined results of the two studies presented in Section V establish a strong and replicable positive association between group cohesion and minimum effort. In this section, we probe the nature of that relationship through three sets of additional analyses.

A. Examining the Predictive Power of Group Cohesion

The first is presented via the three panels of Table 2, where we explore a set of seven models to explain minimum effort using a combination of three independent variables plus a dummy for Study 2 (which reveals the Birmingham subjects to be a little less cooperative than those in Nottingham). Table 2 models are estimated using standard ordered probit with clustering at group-level, since in the experiment groups make multiple decisions.[12] Among the three main explanatory variables, one of them is the group cohesion index (as previously defined), which appears as the sole explanatory variable in Model 1. The regressions pool data for all 8 periods and show that the relationship is stable and strongly significant for Study 1 (Panel A), Study 2 (Panel B) and the combined data set (Panel C). The other six models add two further variables in all possible combinations.

The first is an *homophily index* introduced above in Section II. Including this index in the analysis of Table 2, tests whether group cohesion is simply a proxy for other *observable* characteristics, which vary across groups, in which case we might expect the effects of group cohesion to be weaker in the presence of the homophily index. Consistent with our prior that group cohesion and homophily would be conceptually and empirically related, the results show that the homophily index is a significant predictor of minimum effort in most regressions that exclude group cohesion. Notably, however, it is never significant in regressions that also include group cohesion, while cohesion remains strongly significant in five out of six specifications that include the homophily index (it is weakly significant in the sixth). Hence, the group cohesion index clearly outperforms our best attempt to construct an alternative, homophily-inspired, predictor of minimum effort based on a broad set of tangible group characteristics.

[12] We also account for interdependence of observations across periods by estimating nested random models using GLLAM (see Rabe-Hesketh, Skrondal, and Pickles (2005)). The estimates are reported in the supplemental material, Table SM6.2. 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 conservative standard clustering model is relatively prone to type II errors, while the more powerful nested procedure is relatively prone to type I errors. Our main results, it turns out, are consistent across the two methods.

TABLE 2

ORDERED PROBIT REGRESSIONS OF MINIMUM EFFORT ON GROUP COHESION AND CONTROL VARIABLES

Panel A – Study 1							
Dep. variable: Min. Effort	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Group cohesion	0.448***		0.389***			0.484***	0.452***
	(0.105)		(0.123)			(0.136)	(0.143)
Homophily index				3.871**	2.252	-1.038	-2.032
				(1.809)	(0.233)	(2.315)	(2.292)
Share (principal component)		0.277**	0.164		0.233*		0.184
		(0.116)	(0.129)		(0.126)		(0.130)
Log-likelihood	-644.2	-670.5	-639.7	-681.6	-666.7	-643.6	-634.6
# level 1 (resp. 2) units	520 (65)	520 (65)	520 (65)	520 (65)	520 (65)	520 (65)	520 (65)

Panel B – Study 2: Pre-registered replication independently conducted at the University of Birmingham							
Dep. variable: Min. Effort	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Group cohesion	0.388***		0.300***			0.325**	0.251*
	(0.097)		(0.108)			(0.128)	(0.133)
Homophily index				6.400***	4.645*	2.640	2.254
				(2.468)	(2.559)	(2.883)	(2.845)
Share (principal component)		0.370***	0.221*		0.298***		0.211*
		(0.095)	(0.114)		(0.101)		(0.111)
Log-likelihood	-569.0	-583.1	-557.5	-592.8	-569.2	-565.3	-554.9
# level 1 (resp. 2) units	552 (69)	552 (69)	552 (69)	552 (69)	552 (69)	552 (69)	552 (69)

Panel C – Study 1 and 2 combined							
Dep. variable: Min. Effort	(15)	(16)	(17)	(18)	(19)	(20)	(21)
Group cohesion	0.414***		0.338***			0.391***	0.335***
	(0.074)		(0.083)			(0.095)	(0.100)
Homophily index				5.025***	3.240**	0.777	0.113
				(1.486)	(1.575)	(1.803)	(1.770)
Share (principal component)		0.327***	0.192**		0.269***		0.192**
		(0.073)	(0.085)		(0.079)		(0.084)
Study 2 (dummy variable)	-0.342*	-0.360*	-0.414**	-0.303	-0.398**	-0.351*	-0.415**
	(0.019)	(0.193)	(0.195)	(0.195)	(0.196)	(0.191)	(0.195)
Log-likelihood	-1231.1	-1270.7	-1211.6	-1295.7	-1255.6	-1230.4	-1211.6
# level 1 (resp. 2) units	1072 (134)	1072 (134)	1072 (134)	1072 (134)	1072 (134)	1072 (134)	1072 (134)

NOTES.—Data from Periods 1 to 8. Explanatory variables are at group level. Variable definition and construction are in the supplemental material, section SM3. Period dummies (always included, relative to Period 1) are significantly negative (at $p < 0.05$). Controls for individual effects: clustering. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

A possible interpretation of the relation between cohesion and effort is that the members of high cohesion groups – by virtue of tending to know one another – might have agreed to share their payoffs, a fact which if true would change the payoff structure of the weak-link game, rendering cooperation potentially easier to achieve.[13] While some researchers might view this as an argument for avoiding experiments with real groups to prevent contaminating the lab with awkward features of the world, we see it as a challenge to the worthwhile study of real groups and we respond to that challenge in more than one way. The first is via the use of the "*Share*" variable in the analysis of Table 2. As part of the post-experimental questionnaire, we asked participants whether they planned to share their earnings with other group members and whether their expectation of sharing had affected their behavior in the weak-link game. The share variable in Table 2 is based on the principal component of responses to these two tasks (see supplemental material, section SM4.d, for details). Although the results are patchy, there is some evidence that sharing matters, especially when it enters alone and also across the set of models using pooled data (Panel C). Crucially, however, while the inclusion of share tends to reduce the coefficient on group cohesion, the reductions are modest, and cohesion remains strongly significant in all cases where both variables appear. While these results are reassuring, since we cannot be fully confident that self-reports about sharing are entirely reliable, we treat this as suggestive, rather than conclusive, evidence that the predictive effects of cohesion are unlikely to be substantially explained by sharing agreements among friends. In Section VII we underpin this conclusion more solidly, through an experimental manipulation which tests the sharing account more directly.

B. Assessing the Magnitude of Cohesion-Related Cooperation

Having established a robust and highly significant association between group cohesion and minimum effort, as a second set of analyses, we now assess the *magnitude* of this effect. We explore the predictive power of group cohesion by regressing it on minimum effort in the last period (period 8) to generate the predicted probabilities for each possible level of minimum effort, conditional on different levels of group cohesion using pooled data for Studies 1 and 2 (see supplemental material, Fig. SM2.2, for corresponding analysis separately by study).

[13] While the Nash equilibria are unchanged, the individual marginal cost decreases from 50 points in the standard case to 12.5 (which is equal to 50 divided by the 4 group members) in the case in which group members are "fully sharing" their payoffs; the indicator of the harshness of the game environment *h* (see footnote 6) increases from 0.30 in the standard case to 1.2 in the "full sharing" case, making the latter a much less harsh environment where, potentially, cooperation is easier to achieve.

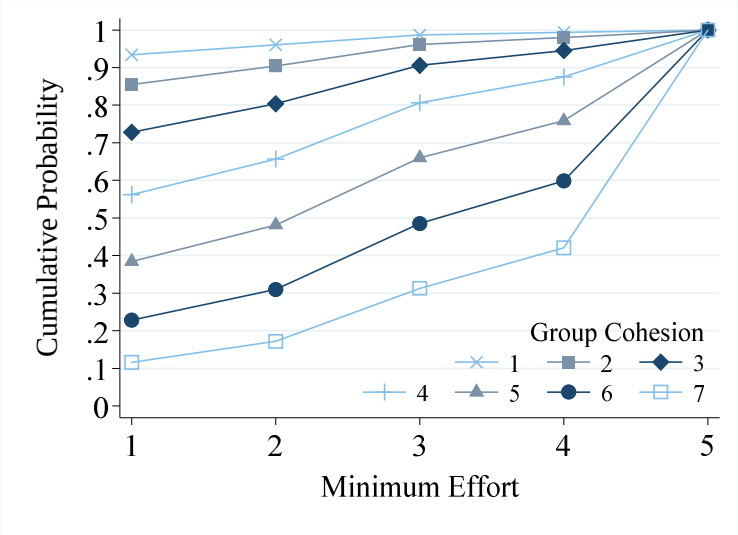


FIG. 5.—Study 1 and 2 combined: predicted CDFs for minimum effort for each level of group cohesion (group-level data from period 8).

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 93 percent). By contrast, a group with maximum possible group cohesion (equal to 7) is unlikely to end up at minimum effort 1 (probability of less than 12 percent) and is predicted to achieve minimum effort of at least 3 with a probability of about 83 percent.

These results point to a powerful effect of group cohesion on minimum effort and, given that, one might wonder how far these results depend on the particular specification of the group cohesion variable? Recall that we calculate group cohesion as the average of the minimum oneness ratings in a group. While this minimum 'envelope' seems a natural statistic, particularly in the context of the weak link game, there is no 'special sauce' involved here: indeed, using the group average of individual oneness ratings as an alternative cohesion metric delivers very similar results to those detailed above (see supplemental material, Table SM6.3).

C. Individual-Level Effort Choice

As a third window on the predictive power of group cohesion, we dig down a level to examine the association between *individual level* effort and group cohesion using pooled data from Study 1 and 2 (see supplemental material Fig. SM2.3, for corresponding analysis separately by study). Fig. 6 shows the distribution of individual effort comparing individuals in groups with

low (panel *a*) and high (panel *b*) group cohesion (these correspond with the two extreme partitions of Figs. 3 and 4). In these panels, for each period, color coding shows the distribution of efforts while the average of individual effort is indicated with a circle.

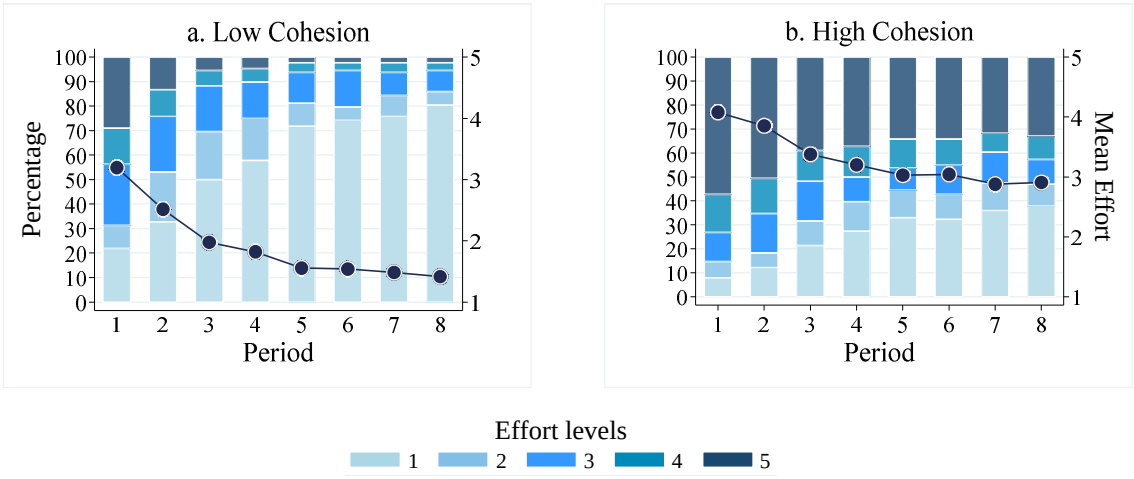


FIG. 6.—Study 1 and 2 combined: distribution of individual efforts over periods. Panel *a*: "Low Cohesion" Partition (32 groups): *group cohesion* ∈ [1, 2]; Panel *b*: "High Cohesion" Partition (41 groups): *group cohesion* ∈ (4, 7] (the relevant panel for the "Medium Cohesion" Partition is reported in the supplemental material Fig. SM2.3). The bars represent the percentage of each effort level ranging from 1 to 5. The y-axes show the relevant percentages. The connected dots represent mean efforts (individual level and measured on the secondary y-axes).

Notice that the time profile of average individual effort is clearly different comparing low and high cohesion groups: for low cohesion groups, it starts just above 3 and descends close to the minimum of 1 by period 8; whereas, for high cohesion groups it starts higher (close to 4) and descends less steeply converging by period 8 to an average effort level of around 3. Persistent differences in the distributions of effort are also apparent comparing low and high cohesion panels (for instance, there is markedly more incidence of efforts above 3 in the high cohesion panels). An econometric analysis also finds a highly significant positive influence of individual average oneness on individual effort choices.[14]

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

[14] A nested random model (GLAMM – see Rabe-Hesketh, Skrondal, and Pickles (2005)) regression shows that individual effort tends to increase with the mean oneness rating of others in their group ($\beta = 0.139$; $p = 0.001$; Study 1 and 2 combined). 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.

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 using pooled data for Studies 1 and 2 strongly confirms this hypothesis.[15]

VII. The Economic Value of Group Cohesion: Financial Incentives and Payoff Sharing

How big is the economic value of the cooperation-enhancing effect of group cohesion we have documented in the previous section? Here, we report the results of two studies (designed to benchmark the cooperation-enhancing effects of group cohesion against those of two other variables: financial incentives ("Bonus Study"), and the likelihood of sharing payoffs ("Share Study") (details in Table SM1.2).

In the Bonus Study, we ran a series of new experimental treatments which varied the bonus (i.e., b in the payoff function π_i of the weak-link game – see Section IV.A) and hence the 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 (see supplemental material, Table SM1.2, for the respective payoff tables). 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 Study (60 per treatment).

The right-most panel of Fig. 7 summarizes the results of the Bonus Study by plotting the average of group minimum effort observed in each Bonus Study treatment. Increasing the bonus has a monotonic impact on the average minimum effort. The effects of the bonus level can be contrasted with the effects of increases in group cohesion by comparison with the left-most panel of Fig. 7 where we plot the expected levels of minimum effort as a function of group cohesion based on the predicted probabilities displayed in Fig. 5 (the "supply curve" of minimum effort as a function of group cohesion). At bonus level 6, the average minimum effort

[15] 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; Study 1 and 2 combined) regressing the indicator variable on the individual mean of the three oneness ratings shows a significant and positive effect ($\beta = 0.102$, $p < 0.001$).

is close to the minimum possible value of one and corresponds with the expected minimum effort associated with low cohesion groups (i.e., a cohesion level of approximately 3). The figure shows that substantial increases in the bonus, beyond those used by Brandts and Cooper, are needed to induce average minimum effort levels comparable to those associated with high cohesion. For example, as Fig. 7 illustrates, a bonus level of 22 in the Bonus Study produces an average minimum effort comparable to that expected from groups with a cohesion level of approximately 6. These results show that the economic value of group cohesion, measured by the material incentives needed to match its effort enhancing effects, is substantial.

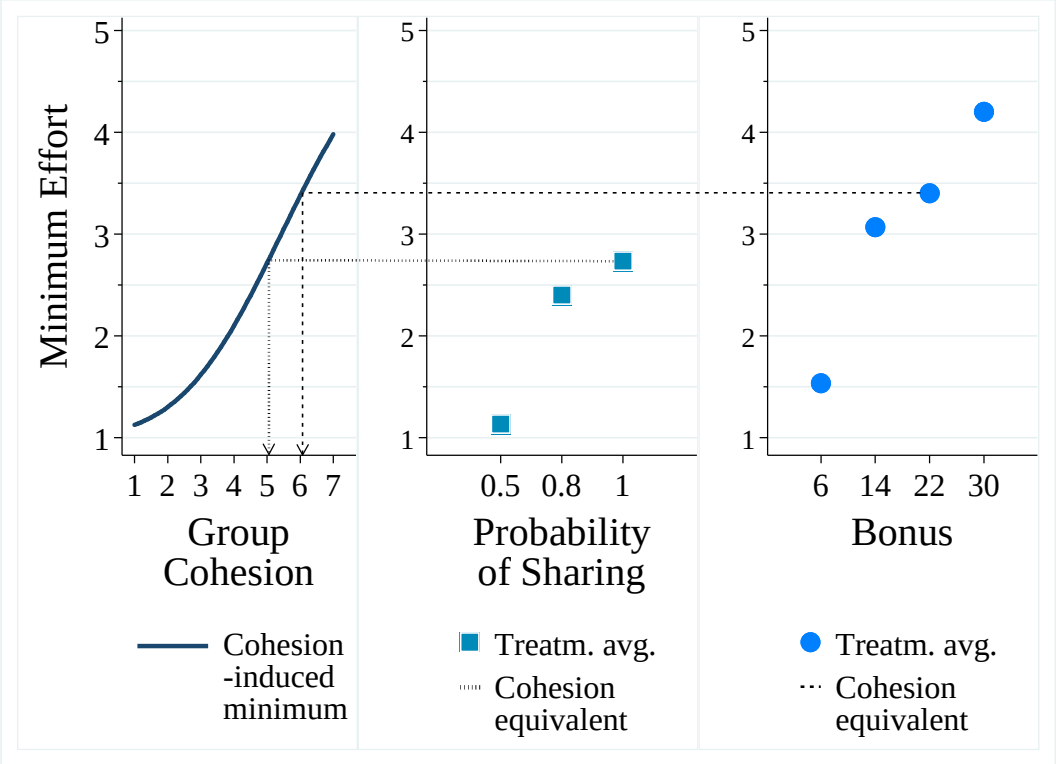


FIG. 7.—The supply of minimum effort as a function of group cohesion in comparison with sharing rules and financial bonuses (group-level data from period 8)*:* In the left panel, the minimum effort on the *y*-axis represents the expected minimum based on the predicted probabilities displayed in Fig. 5, using data from Study 1 and 2 combined. In the central and right panels, minimum effort on the y-axis represents overall treatment averages of minimum efforts. The dotted (dashed) line represents an illustration of a comparison between the effect of group cohesion and sharing rule (financial bonus) by projecting the average minimum effort for the treatment "Probability of Sharing 1" ("Bonus 22") to the cohesion-induced minimum.

As a second benchmarking exercise, we ran the Sharing Study. Like for the Bonus Study, fresh participants were recruited *individually* so that each subject typically did not know any other participant. In the Sharing Study, subjects played the weak-link game of Table 1 (where $b = 6$) following other standard procedures used across our studies (including the Bonus Study) but with the distinguishing feature that, before making their decisions in the weak-link game, subjects were told that there was some probability that we would pool all individual earnings

within each group and share them equally among group members. Part of the interest in running these treatments was hinted at earlier when we noted that a possible interpretation of our main results is that subjects who know one another might plan to share their payoffs, thereby generating a positive association between group cohesion and cooperation. We implemented three versions of this protocol (n = 60 each) with the known probability of sharing being either 0.5, 0.8 or 1. This allows us to assess an upper bound for the impact of sharing (when sharing is certain) and its sensitivity to different levels of uncertainty associated with a potential sharing arrangement.

The results for these three treatments are presented as the middle panel of Fig. 7. The treatment where sharing is certain generates an average minimum effort of 2.73 which is shown on the figure to be comparable to the expected minimum effort associated with a group cohesion of very close to 5. When the likelihood of sharing is only 50%, average minimum effort falls dramatically to an expected minimum level only slightly above 1. While this evidence does not allow us to rule out that expectations of sharing might have played some role in our data, it does count against it being a convincing explanation of the broad patterns in our data. This is so partly because the ceiling of the sharing effect is well below the predicted effect of maximal cohesion (=7) and because some degree of uncertainty – quite likely to be a part of actual sharing arrangements – rapidly diminishes the impact of sharing.

The results of the Sharing Study are interesting for the further reason that, with full sharing of payoffs, a subject is motivated to put equal weight on each group member's earnings including their own. As such, one can interpret that 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). Viewed in this way, the results point to the possibility that social preferences might play some explanatory role. This is an issue that we pick up now via further analysis which investigates both this, and the role played by beliefs.

VIII. Explaining the Power of Group Cohesion: Beliefs and Social Preferences

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 two natural channels through which one might model the impact of group cohesion on individual decisions: via its impacts on *social preferences* and *beliefs*.

Both theoretical and empirical considerations render it plausible to suppose that the preference and belief 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 III and Gächter, Starmer, and Tufano (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) and have common experiences supporting implicit learning (enhanced beliefs). While this interconnectedness means that it will be difficult to separate, cleanly, the impacts of the different channels, we take an econometric approach and use mediation analysis to assess the extent to which the effects of group cohesion on minimum effort can be understood as operating indirectly via effects on beliefs and/or social preferences.

To facilitate this, recall that as part of the setup for Study 2, we included independent measurements of beliefs about the decisions other group members would take and their social preferences towards their group members. Specifically, immediately after entering their effort decision for each round of the weak-link game, but before knowing what others had done, each participant entered their best guess about what would be the minimum effort in that round.[16] Then, at the end of the study, we measured participants' social preferences via a set of standard Social Value Orientation tasks: specifically, we adopted the "Social Value Orientation Slider Measure" by Murphy, Ackermann, and Handgraaf (2011) consisting of 15 allocation tasks to measure subjects' social preferences toward each group member. In this approach, each allocation task is akin to a dictator game. Tasks differ in the allocations available and allow measurement of altruistic, pro-social, individualistic, competitive and spiteful preferences.[17] We calculated group-level beliefs as the group average of the individual beliefs per group and

[16] In line with Schlag, Tremewan, and van der Weele (2015), p. 484, we opted for a non-incentivized belief elicitation because ours were fresh subjects with no clear incentive to misreport, facing a straightforward elicitation task embedded in a multi-task experiment in which hedging could otherwise have been a problem. See the supplemental material, section SM4.c, for further details.

[17] In each group, members were randomly allocated to roles A, B, C and D, respectively. Each member was matched with a recipient (from their own group) to whom they could allocate resources. Specifically, participant A could give resources to B, B to C, C to D and D could give to A. Each participant knew which recipient they were matched with, but the recipient did not know who could give them resources. See the supplemental material, section SM4.c, for further details.

the group-level social preference as the group average of the individual social value orientations per group.

We use responses to these two sets of tasks as key inputs to a simultaneous equation model, estimated via the following three linear equations using group-level data:

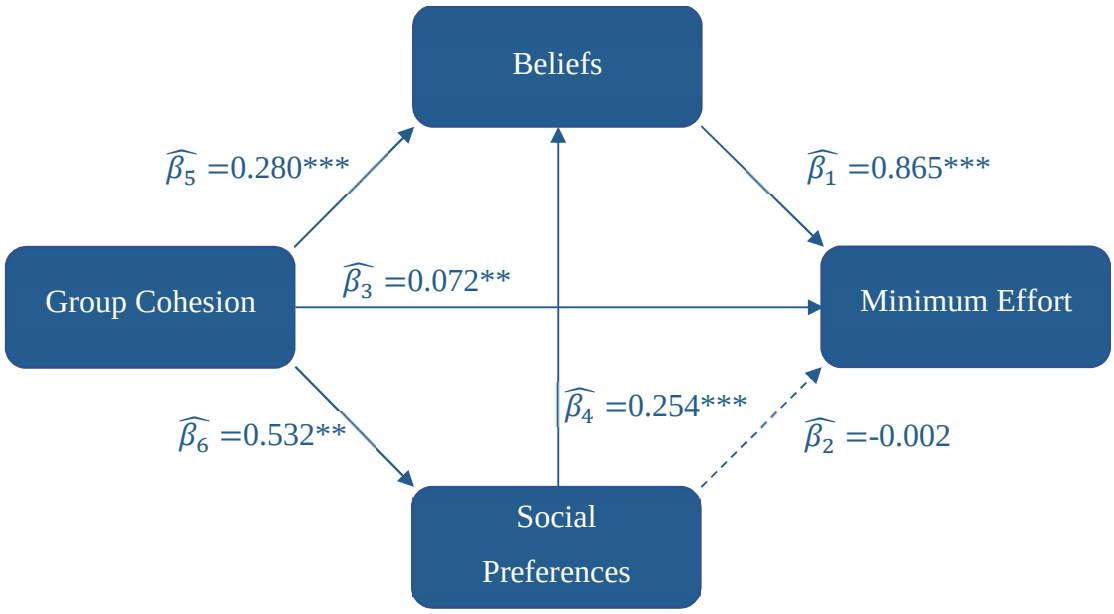
$$Minimum\_Effort = \alpha_1 + \beta_1\, Beliefs + \beta_2\, Social\_Preferences + \beta_3\, Group\_Cohesion + \varepsilon_1 \quad (1)$$

$$Beliefs = \alpha_2 + \beta_4\, Social\_Preferences + \beta_5\, Group\_Cohesion + \varepsilon_2 \quad (2)$$

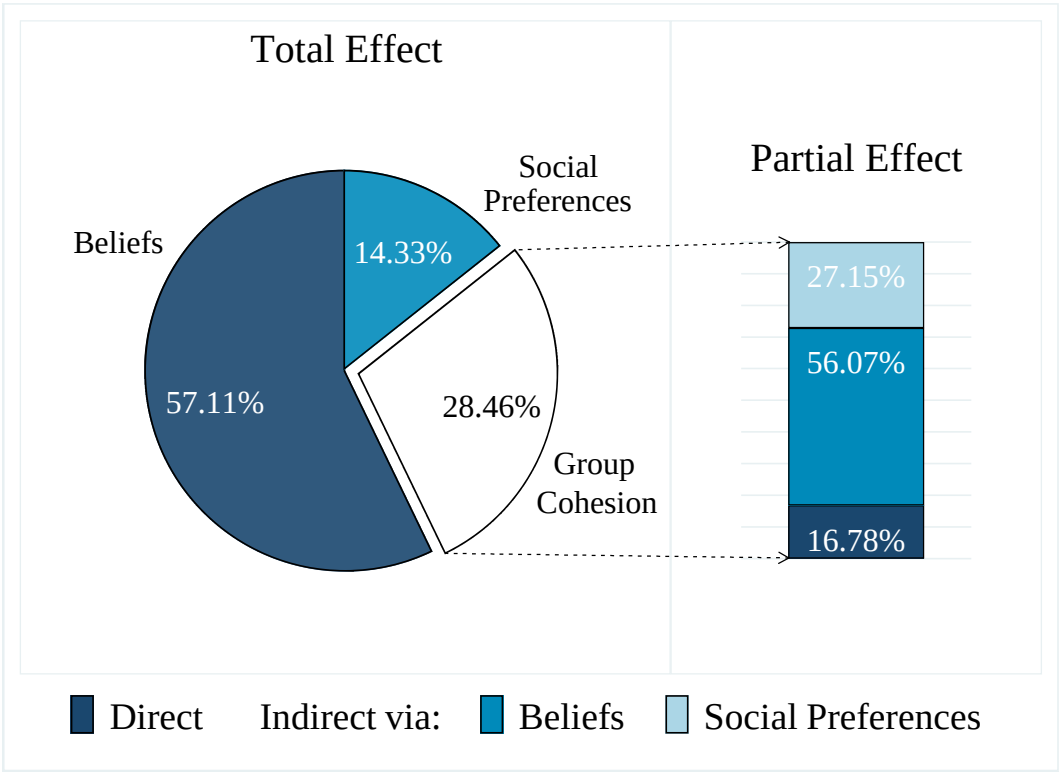
$$Social\_Preferences = \alpha_3 + \beta_6\, Group\_Cohesion + \varepsilon_3 \quad (3)$$

The model is inspired by the rational choice perspective in positing beliefs and (social) preferences as proximate determinants of minimum effort (as per Eq. 1). The model is also constructed to embody two pathways motivated and conjectured above: in the model, both beliefs (Eq. 2) and social preferences (Eq. 3) may be influenced by group cohesion. The full logic of the model is most easily illustrated, graphically, via the top panel of Fig. 8. Group cohesion is treated as the unique (a priori) exogenous variable which can in principle influence minimum effort directly, and indirectly, via preference or belief channels. In the spirit of models linking social preferences and beliefs (e.g., Dufwenberg, Gächter, and Hennig-Schmidt (2011)), the model also allows social preferences to influence beliefs. The *direct* effects of beliefs and social preferences are represented, respectively, by β_1 and β_2, while coefficient β_3 captures the direct effect of group cohesion on minimum effort. β_5 and β_6 capture the influences of group cohesion on beliefs and social preferences, respectively, while β_4 captures the impact of social preferences on beliefs. Although essentially very simple, the model provides a tool for assessing the relative importance of beliefs and social preferences as channels mediating the impact of group cohesion on effort choices.[18]

[18] The general approach we adopt here is very similar in spirit to the mediation analysis reported in Kosse et al. (forthcoming).



a. Mediation Analysis



b. Decomposition of Total/Partial Effect on Minimum Effort

FIG. 8.—Study 2: modelling how group cohesion affects minimum effort. Panel *a* reports econometric estimates of the simultaneous linear equations 1-3 above (all coefficients are standardized; clustering at group level applied). In panel *a*, solid arrows represent detected statistically significant effects, while the dashed arrows represents the absence of any detected statistically significant effect. Panel *b* reports the decomposition of the total/partial effect on minimum effort based on estimates shown in panel *a*.

Estimated parameters are included on the panel adjacent to their relevant branch of the diagram and these indicate a statistically significant positive effect along every branch except for one: there is no significant direct effect from social preferences to minimum effort (this link is marked with a dashed arrow). Hence, we find that group cohesion impacts minimum effort directly and through its impacts on both beliefs and social preferences, though the last of these channels works entirely through the secondary effect of social preferences on beliefs. Detailed estimation results are in the supplemental material, Table SM6.4.

Having identified three 'active' channels through which group cohesion influences minimum effort, we then perform a decomposition analysis to assess the relative contributions of these different channels (similar to, e.g., Kosse et al. (forthcoming)). While the pie chart in panel *b* provides a summary of the complete decomposition analysis for the whole model, for the purposes of assessing the relative impact of the channels through which group cohesion impacts minimum effort, our primary interest lies in examining the relative sizes of the stripes in the bar chart of panel *b* (these decompose the total effect of group cohesion into its three constituent paths). The path that goes from group cohesion through beliefs accounts for 56.07% of the total effect of group cohesion on minimum effort. While the indirect path through social preferences also accounts for a significant proportion of the total effect (27.15%), this path too operates via the belief channel suggesting that social preferences play a subsidiary role in terms of both scale and, in the absence of a direct effect, mechanism. Finally, the direct effect from group cohesion to minimum effort accounts for only 16.78% of the total effect of group cohesion. We interpret the small size of this direct effect as good news because it means that most of the impact of group cohesion in our data can be readily explained through its influence on the familiar rational choice concepts of beliefs and preferences. The fact that the lion's share of the work is done by beliefs stands in sharp contrast to results based on experiments using artificially-induced groups (see Chen and Chen (2011)) and suggests new directions to be considered in theoretical research.

IX. Conclusions

It is hard to deny that social relationships affect many variables that naturally interest economists. An open question is how much they matter and whether economic analysis could 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 extremely well with more detailed measures of personal relationships (Gächter, Starmer, and Tufano (2015)). We used our measure of group cohesion, which is a simple group-level statistic of the oneness scale, to study the cohesion of real groups. We showed that group cohesion varies across groups as predicted by relevant sociological and psychological literature and is stable based on test-retest measurement.

Using an extensive set of experiments involving over 1200 participants and including a variety of robustness tests, benchmarking exercises, and an independent pre-registered replication, we examined the predictive power of 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, high cohesion groups do much better in terms of the equilibria they achieve in weak-link games, and sufficiently high cohesion appears a necessary condition for cooperation. 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.

We used an econometric approach to explore possible mechanisms underpinning the impact of group cohesion on cooperation and found that group cohesion shapes both beliefs and social preferences but with beliefs emerging as the primary channel.

Our results also have significance in the context of organizational performance (e.g., Akerlof and Kranton (2005); Ashraf and Bandiera (2018)). If group cohesion is a significant determinant of desirable team or group outcomes across a wide range of organizational settings, then having a good tool to quantify it will facilitate a wide range of productive applied research. And, for those with interests in engineering better organizational 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 organizations already invest.

More generally, our paper reveals the importance of real human relationships as economically significant factors of team production. The scale and replicability of these effects provide motivation for new streams of research to explore broader consequences of group cohesion exploiting the conceptual foundations and practical tools provided by this paper.

REFERENCES

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

Anderson, Simon P., Jacob K. Goeree, and Charles A. Holt. 2001. “Minimum-Effort Coordination Games: Stochastic Potential and Logit Equilibrium." *Games and Economic Behavior* 34 (2): 177-99.

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

Ashraf, Nava, and Oriana Bandiera. 2018. “Social Incentives in Organizations." *Annual Review of Economics* 10 (1): 439-63.

Baccara, Mariagiovanna, and Leeat Yariv. 2013. “Homophily in Peer Groups." *American Economic Journal: Microeconomics* 5 (3): 69-96.

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

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

Bohnet, Iris, and Bruno S. Frey. 1999. “Social Distance and Other-Regarding Behavior in Dictator Games: Comment." *American Economic Review* 89 (1): 335-39.

Brañas-Garza, Pablo, Ramón Cobo-Reyes, María Paz Espinosa, Natalia Jiménez, Jaromír Kovářík, and Giovanni Ponti. 2010. “Altruism and Social Integration." *Games and Economic Behavior* 69 (2): 249-57.

Brandts, Jordi, and David J. Cooper. 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.

———. 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. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton: Princeton University Press, 2003.

Camerer, Colin F., Anna Dreber, and Magnus Johannesson. "Replication and Other Practices for Improving Scientific Quality in Experimental Economics." In *Handbook of Research Methods and Applications in Experimental Economics*, edited by Arthur Schram and Aljaž Ule, 83-102. Cheltenham: Edward Elgar Publishing, 2019.

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

Charness, Gary, and Uri Gneezy. 2008. “What's in a Name? Anonymity and Social Distance in Dictator and Ultimatum Games." *Journal of Economic Behavior & Organization* 68 (1): 29-35.

Charness, Gary, Uri Gneezy, and Michael A. Kuhn. 2013. “Experimental Methods: Extra-Laboratory Experiments-Extending the Reach of Experimental Economics." *Journal of Economic Behavior & Organization* 91: 93-100.

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

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

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

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

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

Dufwenberg, Martin, Simon Gächter, and Heike Hennig-Schmidt. 2011. “The Framing of Games and the Psychology of Play." *Games and Economic Behavior* 73 (2): 459-78.

Engelmann, Dirk, and Hans-Theo Normann. 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, Chris Starmer, and Fabio Tufano. 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 John Roberts, eds. *The Handbook of Organizational Economics*. Princeton and Oxford: Princeton University Press, 2013.

Goette, Lorenz, David Huffman, and Stephan Meier. 2012. “The Impact of Social Ties on Group Interactions: Evidence from Minimal Groups and Randomly Assigned Real Groups." *American Economic Journal: Microeconomics* 4 (1): 101-15.

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

Guala, Francesco, and Antonio Filippin. 2017. “The Effect of Group Identity on Distributive Choice: Social Preference or Heuristic?" *Economic Journal* 127 (602): 1047–68.

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

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

Harrison, Glenn W., and John A. List. 2004. “Field Experiments." *Journal of Economic Literature* 42 (4): 1009-55.

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

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

Kets, Willemien, and Alvaro Sandroni. 2019. “A Belief-Based Theory of Homophily." *Games and Economic Behavior* 115: 410-35.

Knez, M., and Colin F. Camerer. 1994. “Creating Expectational Assets in the Laboratory - Coordination in Weakest-Link Games." *Strategic Management Journal* 15: 101-19.

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

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk. forthcoming. “The Formation of Prosociality: Causal Evidence on the Role of Social Environment." *Journal of Political Economy*.

Leider, Stephen, Markus M. Möbius, Tanya Rosenblat, and Quoc-Anh Do. 2009. “Directed Altruism and Enforced Reciprocity in Social Networks." *Quarterly Journal of Economics* 124 (4): 1815-51.

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

Murphy, Ryan O., and Kurt A. Ackermann. 2014. “Social Value Orientation: Theoretical and Measurement Issues in the Study of Social Preferences." *Personality and Social Psychology Review* 18 (1): 13-41.

Murphy, Ryan O., Kurt A. Ackermann, and Michel J. J. Handgraaf. 2011. “Measuring Social Value Orientation." *Judgment and Decision Making* 6 (8): 771-81.

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

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

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

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

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

Rubin, Zick. 1970. “Measurement of Romantic Love." *Journal of Personality and Social Psychology* 16 (2): 265-+.

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

Schlag, Karl H, James Tremewan, and Joël J van der Weele. 2015. “A Penny for Your Thoughts: A Survey of Methods for Eliciting Beliefs." *Experimental Economics* 18 (3): 457-90.

Starzyk, K. B., R. R. Holden, L. R. Fabrigar, and T. K. MacDonald. 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.

Van Huyck, John B., Raymond C. Battalio, and Richard O. Beil. 1991. “Strategic Uncertainty, Equilibrium Selection, and Coordination Failure in Average Opinion Games." *The Quarterly Journal of Economics* 106 (3): 885-910.

———. 1990. “Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure." *American Economic Review*. 80 (1): 234-48.

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

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

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

> For Online Publication <

Supplemental Material (SM) for

The Surprising Capacity of the Company You Keep: Revealing Group Cohesion as a Powerful Factor of Team Production

Simon Gächter, Chris Starmer and Fabio Tufano*

*University of Nottingham***

December 20, 2019

Table of Contents

SM1: Structure of the Program of Experiments

TABLE SM1.1

EXPERIMENTAL STUDIES

		Research Objectives	Location	Incentive Structure	Recruitment	Allocation	N (Total 1204)
I.	Study 1	Predictive Power of Group Cohesion	CeDEx Lab Nottingham	b=6	Friends	Random (F- or N-matching)	260
II.	Two-week Study	Construct reliability (Test-retest reliability; Task sequencing)	CeDEx Lab Nottingham	b=6	Friends	Random (F- or N-matching)	172
III.	Study 2	Replicating Study 1; Mediational channels: Beliefs, Social Prefs.	BEEL Lab Birmingham	b=6	Friends	Random (F- or N-matching)	276
IV.	50-period Study	Long horizon	CeDEx Lab Nottingham	b=6	Strangers	Random to groups	32
V.	JOTP Study	Categorical group identity; Common culture	JOTP Bern	b=6	JOTP officers	Random to groups	44
VI.	Share Study	To compare the cooperation enhancing effects of group cohesion with sharing rules	CeDEx Lab Nottingham	b=6, Pr(Sharing)=0.5	Strangers	Random to groups	60
				b=6, Pr{Sharing)=0.8	Strangers	Random to groups	60
				b=6, Pr(Sharing)=1	Strangers	Random to groups	60
VII.	Bonus Study	To compare the cooperation enhancing effects of group cohesion with financial bonuses	CeDEx Lab Nottingham	b=6	Strangers	Random to groups	60
				b=14	Strangers	Random to groups	60
				b=22	Strangers	Random to groups	60
				b=30	Strangers	Random to groups	60

NOTE.—The Studies listed above were conducted across three locations: (*i.*) Centre for Decision Research and Experimental Economics (CeDEx) Lab, University of Nottingham, UK; (*ii.*) The Birmingham Experimental Economics Laboratory (BEEL), University of Birmingham, UK; (*iii.*) Junior Officer Training Program (JOTP), Swiss Army, Bern, CH. Study 2 was a pre-registered (at The American Economic Association's Registry for Randomized Controlled Trials – RCT ID: AEARCTR-0003566) replication independently conducted at the BEEL lab by in-house experimenters.

TABLE SM1.2

PAYOFF MATRICES FOR THE WEAK-LINK GAMES WITH VARIOUS BONUS LEVELS

Bonus 6						*h=0.30*
		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

Bonus 14						*h=0.70*
		Minimum Effort				
		1	2	3	4	5
Effort by Player *i*	1	280				
	2	230	370			
	3	180	320	460		
	4	130	270	410	550	
	5	80	220	360	500	640

Bonus 22						*h=1.10*
		Minimum Effort				
		1	2	3	4	5
Effort by Player *i*	1	360				
	2	310	530			
	3	260	480	700		
	4	210	430	650	870	
	5	160	380	600	820	1040

Bonus 30						*h=1.50*
		Minimum Effort				
		1	2	3	4	5
Effort by Player *i*	1	440				
	2	390	690			
	3	340	640	940		
	4	290	590	890	1190	
	5	240	540	840	1140	1440

NOTE.—The payoffs are expressed in points. *h* is an indicator of the harshness of the game environment (see footnote 6 in the paper for further details).

SM2: Study 1 and 2: Additional Descriptive Analysis

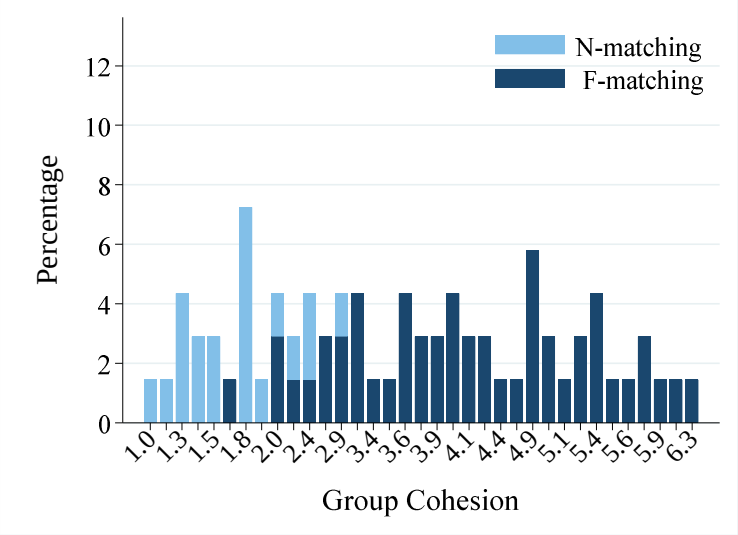


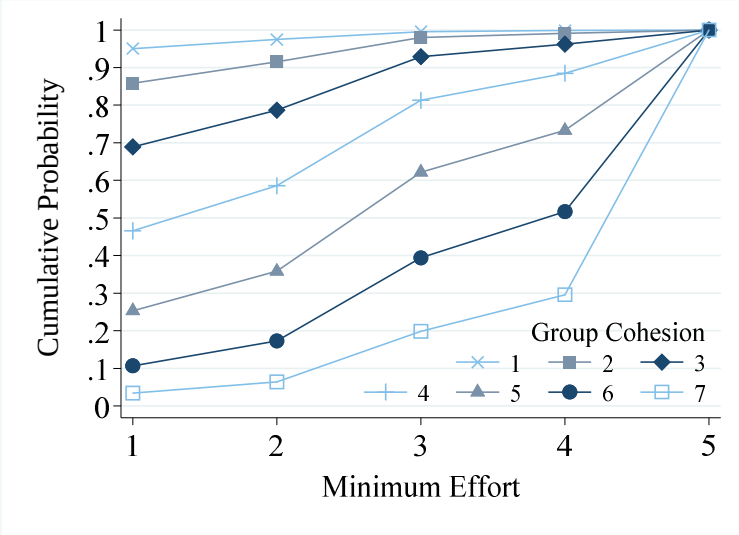
FIG. SM2.1.—Study 2: The distribution of group cohesion under F- and N-matching. Note: in Study 2 we have groups scoring cohesion levels beyond the 5.5 maximum of Study 1.

TABLE SM2.1

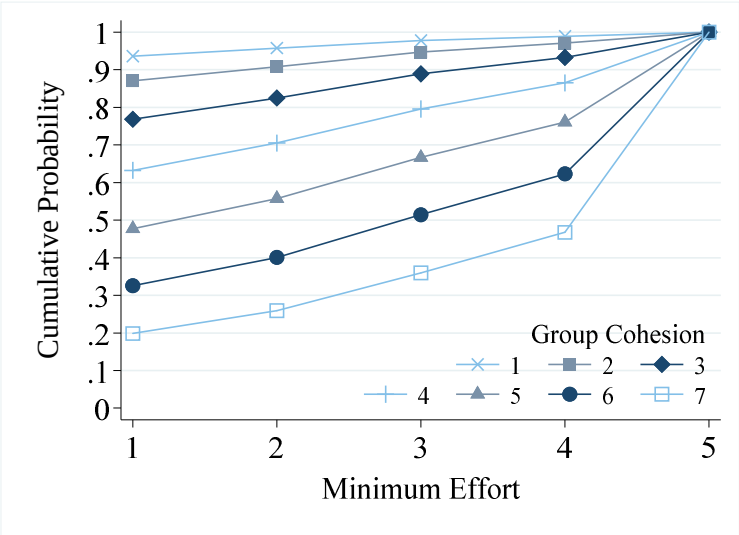
STUDY 1 AND 2: TESTING FOR EQUALITY OF VARIANCE WITHIN AND BETWEEN GROUPS

#	Individual-level Variable	Study 1 ANOVA F	Study 1 ANOVA R^2	Study 1 K.-W. test χ^2 (with ties)	Study 2 ANOVA F	Study 2 ANOVA R^2	Study 2 K.-W. test χ^2 (with ties)
1	Nationality^i	26.700***	0.897	167.807***	22.950***	0.882	170.262***
2	Age^i	8.300***	0.730	155.573***	14.700***	0.828	161.400***
3	Field of studies^i	4.780***	0.609	114.082***	4.050***	0.570	113.635***
4	Gender^i	4.610***	0.601	112.323***	2.930***	0.489	95.344***
5	Team perception index^i	2.900***	0.486	92.868***	4.30***	0.584	108.537***
6	Religiousness^i	2.630***	0.462	82.772***	4.28***	0.583	109.375***
7	No. of club memberships^i	2.350***	0.434	80.388***	1.920***	0.385	69.946**
8	Prior group interactions^i	3.870***	0.558	77.187***	4.140***	0.575	113.732***
9	Share intention^i	2.230***	0.421	76.650***	1.870***	0.379	73.572**
10	Share influence^i	1.960***	0.390	76.527***	2.360***	0.436	85.416***
11	Monthly budget^i	1.056**	0.337	74.820***	2.390***	0.439	92.216***
12	No. of siblings^i	2.020***	0.397	71.930***	2.160***	0.413	88.313***
13	Cohabitee^i	2.190***	0.417	69.942**	1.810***	0.371	81.649***
14	City size^i	1.750***	0.364	68.677**	1.520**	0.332	65.152*
15	Trust index^i	1.640**	0.348	66.685**	1.19	0.216	56.398
16	Risk attitude^i	1.730***	0.361	64.846**		*n.a.*	
17	Self-financed^i	1.550**	0.336	63.338**	1.290	0.297	60.666
18	Political attitude	1.360**	0.308	60.991*	1.970***	0.392	81.244***
19	Empathy index^i	1.450**	0.321	60.298*	1.290	0.297	53.831
20	Income rank^i	1.280	0.295	60.255*	1.830***	0.374	69.413**
21	Behavioral risk attitude^i	1.580**	0.376	59.573*	1.420*	0.317	59.522
22	Trust strangers^i	1.140	0.271	50.704	1.530**	0.333	65.021*
23	Future happiness^i	1.020	0.249	49.863	1.610**	0.344	61.682*
24	Current happiness^i	1.060	0.257	47.546	1.380*	0.311	63.135
25	Loss attitude^i	1.000	0.260	44.501		*n.a.*	

NOTE.—F-matching data only. Individual-level variables are ordered according to scores in the Kruskal-Wallis test (denoted "K.-W." in the table) in Study 1. Variable specification is given below in section SM3. ***, ** and * denote significance at the 1, 5 and 10 percent levels respectively. Risk attitude^i and Loss attitude^i were not elicited in Study 2.



a. Study 1



b. Study 2: Pre-registered replication independently conducted at the University of Birmingham

FIG. SM2.2.—Predicted CDFs for minimum effort for each level of group cohesion (group-level data from period 8).

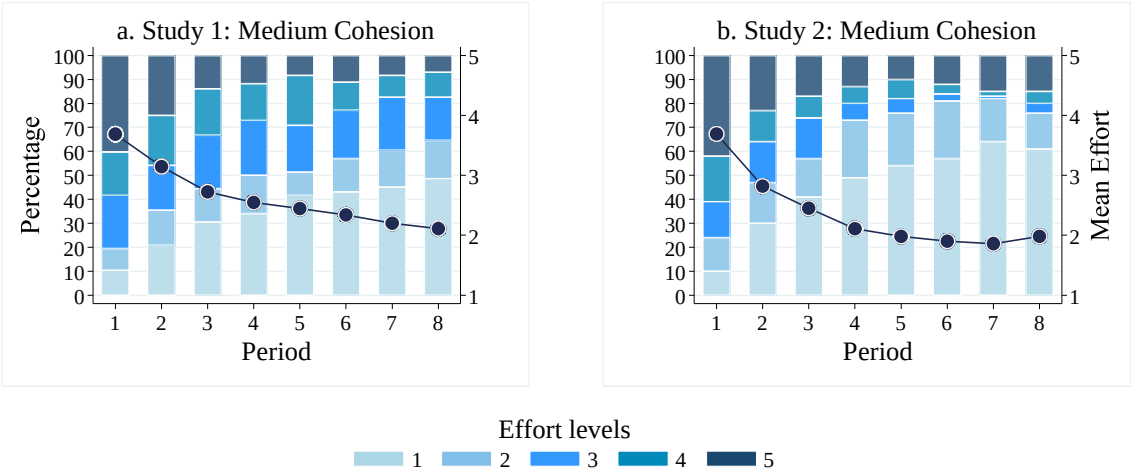


FIG. SM2.3.—Distribution of individual efforts over periods in Study 1 and 2 for the "Medium Cohesion" partition. The "Medium Cohesion" partition has 36 groups in Study 1 and 25 groups in Study 2, respectively. The bars represent the percentage of each effort level ranging from 1 to 5. The y-axes show the relevant percentage. The connected dots represent mean efforts (individual level and measured on the secondary y-axes).

SM3: Individual and group characteristics

SM3.a. 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 A3. (The text in squared brackets reports the reference number of the relevant question in section SM2.d).

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

22. *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.*]

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

SM3.b. Group-level Variable Construction

The group-level variables have labels corresponding with the relevant individual-level variables (see above subsection SM3.a) from which we construct them, but the superscript *i* has been dropped. The group-level variables are constructed as follows. *Prior group interactions*, *Empathy index*, *Share intention*, *Share influence*, *Team perception index*, *Trust index* and *Trust strangers* are constructed by taking the group means of the corresponding individual-level variables. *Share (principal component)*

is obtained by taking the principal component of *Share intention* and *Share influence*. *Trust (principal component)* is calculating by extracting the principal component from *Trust index* and *Trust strangers*. We describe the construction of the group-level variable *Homophily index* in the next subsection.

SM3.c. Homophily Index Construction

To construct the overall homophily index at group level, we first generate group-level homophily indices for each individual-level variable listed in section SM3.a above, with the exception of *Prior group interactions*^{i}, *Empathy index*^{i}, *Team perception index*^{i}, *Trust index*^{i} and *Trust strangers*^{i}.[1] The homophily indices have labels corresponding with the relevant individual-level variables from which we construct them, but the superscript *i* has been dropped. The group-level homophily indices for each individual-level variable are constructed as follows.

1. *Gender homophily index*: From the two mutually-exclusive categories (*i.*) Female; (*ii.*) Male, calculate the highest proportion in each group of members from one of the two aforementioned categories.
2. *Age homophily index*: First, generate 4 mutually-exclusive categories: (*i.*) Under 20; (*ii.*) Under 24; (*iii.*) Under 29; (*iv.*) Over 30. Second, calculate the highest proportion in each group of members from one of the four above categories.
3. *Field study homophily index*: From the five mutually-exclusive categories (*i.*) arts, (*ii.*) business and economics, (*iii.*) social science, (*iv.*) medicine and (*v.*) science, calculate the highest proportion in each group of members from one of the five above categories.
4. *Nationality homophily index*: From the 3 mutually-exclusive categories (*i.*) British, (*ii.*) EU and (*iii.*) Other, calculate the highest proportion in each group of members from one of the three aforementioned categories.
5. *No. of siblings homophily index*: First, generate 3 mutually-exclusive categories: (*i.*) no siblings; (*ii.*) one or two siblings; (*iii.*) three or more siblings. Second, calculate the highest proportion in each group of members from one of the three above categories.

[1] The group-level variables constructed on the basis respectively of *Prior group interactions*^{i} and *Empathy index*^{i} figure as controls in Table 1 in the main paper and Table A2 in the appendix. *Team perception index*^{i} forms the basis for generating the group-level variable *Team perception index* while the group-level variable *Trust (principal component)* corresponds to the principal component of *Trust index*^{i} and *Trust strangers*^{i}. *Team perception index* and *Trust (principal component)* appear as independent explanatory variables in Table SM1.1 above. Details on the construction of these group-level variables are provided above in subsection SM3.b.

6. *Income rank homophily index*: First, generate 3 mutually-exclusive categories: (*i.*) far below and below average; (*ii.*) average; (*iii.*) above and far above average. Second, calculate the highest proportion in each group of members from one of the three above categories.
7. *City size homophily index*: First, generate 3 mutually-exclusive categories: (*i.*) up to 10,000 inhabitants (this is obtained by merging the "up to 2,000 inhabitants" with the "2,000 to 10,000 inhabitants"); (*ii.*) 10,000 to 100,000 inhabitants; (*iii.*) more than 100,000 inhabitants. Second, calculate the highest proportion in each group of members from one of the three above categories.
8. *Cohabitee homophily index*: First, generate 3 mutually-exclusive categories: (*i.*) up to 3 cohabitees; (*ii.*) 4 or 5 cohabitees; (*iii.*) more than 5 cohabitees. Second, calculate the highest proportion in each group of members from one of the three above categories.
9. *Monthly budget homophily index*: First, generate 3 mutually-exclusive categories: (*i.*) lower than £200; (*ii.*) between £200 and £300; (*iii.*) more than £300. Second, calculate the highest proportion in each group of members from one of the three above categories.
10. *Self-financed homophily index*: First, generate 3 mutually-exclusive categories: (*i.*) up to 33%; (ii.) between 33% and 66%; (iii.) more than 66%. Second, calculate the highest proportion in each group of members from one of the three above categories.
11. *No. of memberships homophily Index*: First, generate 3 mutually-exclusive categories: (*i.*) up to 2; (*ii.*) between 2 and 4; (*iii.*) more than 4. Second, calculate the highest proportion in each group of members from one of the three above categories.
12. *Religiousness homophily index*: First, generate 3 mutually-exclusive categories: (*i.*) up to 2 on a Likert scale from 1 "Not at all religious" to 7 "Very religious"; (*ii.*) between 2 and 4 on a Likert scale from 1 "Not at all religious" to 7 "Very religious"; (*iii.*) more than 4 on a Likert scale from 1 "Not at all religious" to 7 "Very religious". Second, calculate the highest proportion in each group of members from one of the three above categories.
13. *Political attitude homophily index*: First, generate 3 mutually-exclusive categories: (*i.*) up to 2 on a Likert scale from 1 "Left" to 7 "Right"; (*ii.*) between 2 and 4 on a Likert scale from 1 "Left" to 7 "Right"; (*iii.*) more than 4 on a Likert scale from 1 "Left" to 7 "Right". Second, calculate the highest proportion in each group of members from one of the three above categories.
14. *Current happiness homophily index*: First, generate 3 mutually-exclusive categories: (*i.*) up to 4 (included) on a Likert scale from 1 "Not at all satisfied" to 10 "Absolutely satisfied"; (*ii.*) between 5 and 7 a Likert scale from 1 "Not at all satisfied" to 10 "Absolutely satisfied"; (*iii.*) between 8 and 10 on Likert scale from 1 "Not at all satisfied" to 10 "Absolutely satisfied". Second, calculate the highest proportion in each group of members from one of the three above categories.

15. *Future happiness homophily index*: First, generate 3 mutually-exclusive categories: (*i.*) up to 4 (included) on a Likert scale from 1 "Not at all satisfied" to 10 "Absolutely satisfied"; (*ii.*) between 5 and 7 a Likert scale from 1 "Not at all satisfied" to 10 "Absolutely satisfied"; (*iii.*) between 8 and 10 on Likert scale from 1 "Not at all satisfied" to 10 "Absolutely satisfied". Second, calculate the highest proportion in each group of members from one of the three above categories.

We construct the *Homophily index* as the overall average of the above fifteen homophily indices.

SM4: Experimental Procedures and Instructions

Section SM4.a reports the invitation letter for the Study-2 experiment. Section SM4.b documents the wording used for the welcome of Study-2 experimental participants and the initial oral instructions. Section SM4.c presents the Study-2 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 SM4.d details the Study-2 post-experimental control tasks (including related summary statistics). Study-2 experimental procedures and instructions were identical to those used in Study 1 with the exception of minor changes necessary to accommodate location-specific features and the fact that in Study 1 participants' risk (à la Holt and Laury, 2002) and loss (à la Gächter et al., 2010) attitudes were elicited where in Study 2 participants' game beliefs and social value orientations were elicited instead.

SM4.a. Invitation Letter

Dear #Name# #Surname#,
You registered with the Birmingham Experimental Economics Lab (BEEL) 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 60 minutes, and at the end YOU and

your THREE friends will be paid in private and in cash.
The experiment will take place in the Birmingham Experimental Economics lab (BEEL), room 101, University House (aka Business School, O3 on Campus Map) Edgbaston Park Rd, Edgbaston, Birmingham B15 2TY 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 email to communicate the full name and the University of Birmingham email addresses 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 you cannot click on the link, copy it to the clipboard by selecting it, right-click and choosing "Copy", and then paste it into the address line in your browser by right clicking there and choosing "Paste".)
Please be on time and remember to bring your Student ID to the session.
This experiment is named #Experiment_Name#. Please include "#Experiment_Name#" in the subject field of any email you send BEEL regarding this experiment.
Best regards,
The BEEL Team

SM4.b. Welcome and Oral Instructions

Welcome to the Birmingham Experimental Economics Lab!
Thank you very much for participating.
This is an experiment in decision making and it is funded by various Institutions. 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: 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?

SM4.c. 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

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

****

Fig. SM4.1 below presents the screenshot of the "Instructions for Task 1."

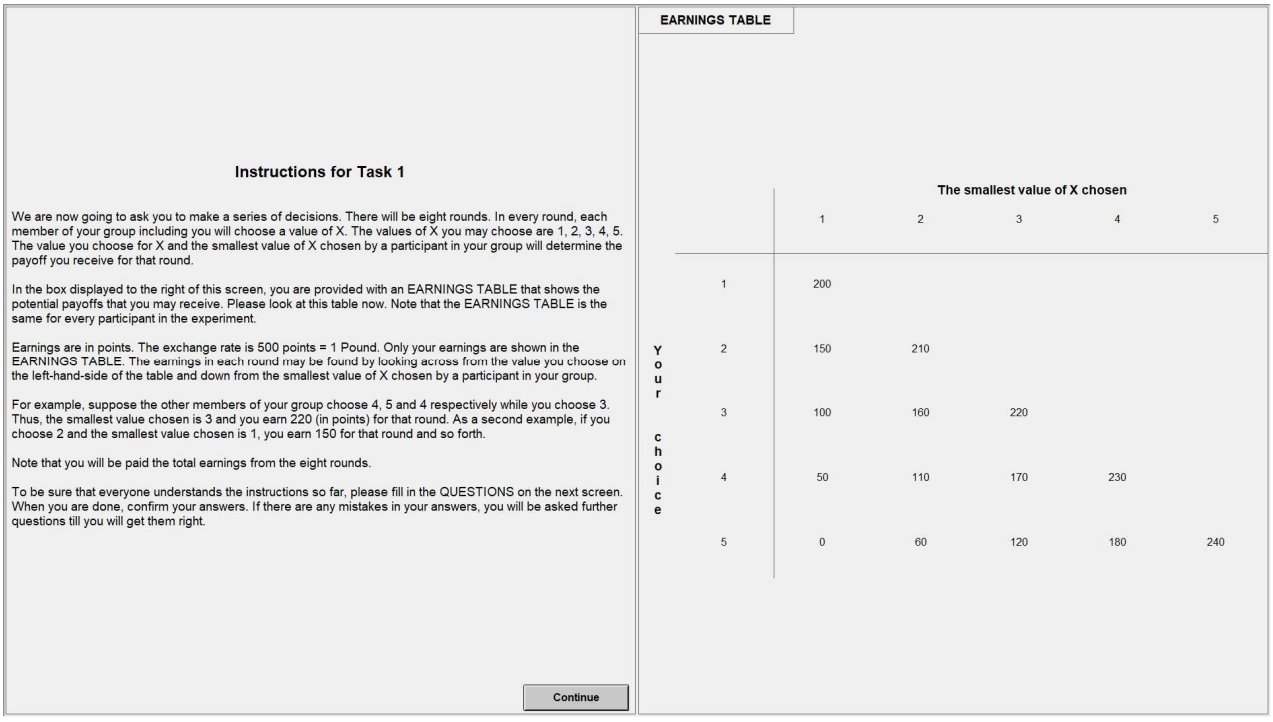


FIG. SM4.1.—Screenshot of "Instruction for Task 1".

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

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

****

Fig. SM4.2 below is the screenshot of the input screen for the first period of the weak-link game.

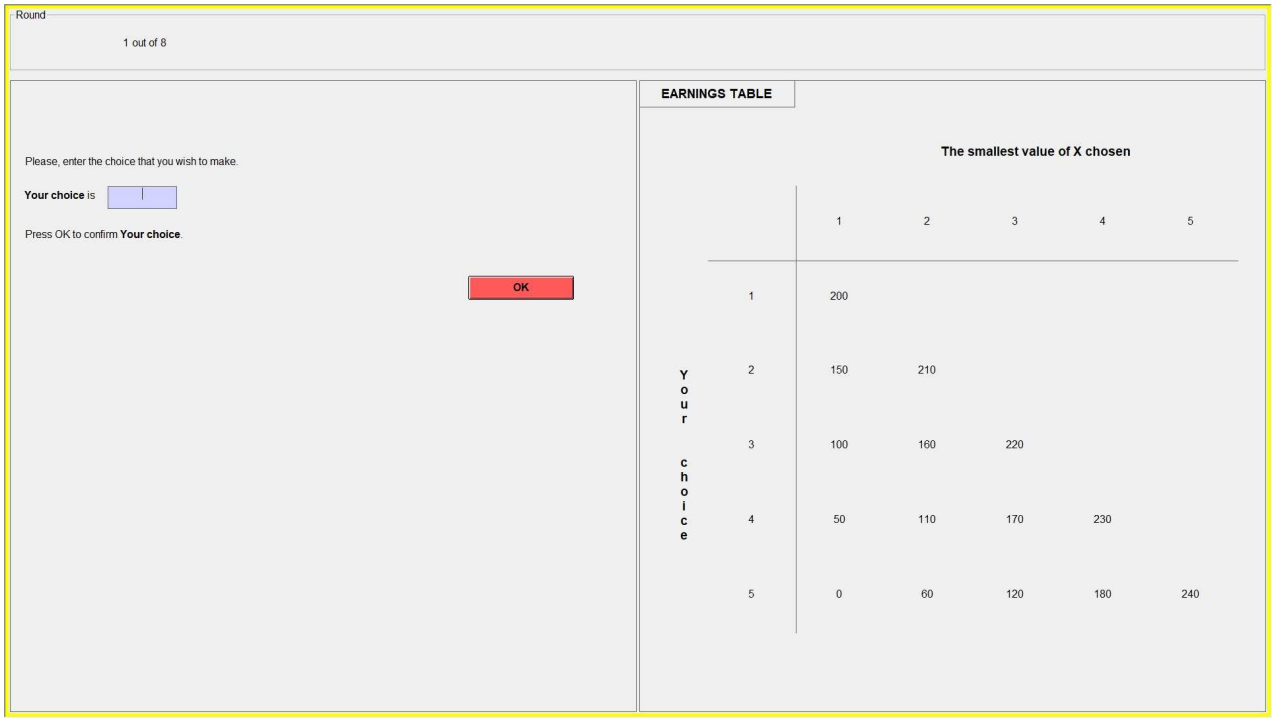


FIG. SM4.2.—Period 1 of the weak-link game: input screen.

In each period, after each participant had entered their chosen effort their computer screen would instruct them to enter their best guess about the smallest value of X and state their confidence. The relevant text for the first period of the game is reported as follows:

****

Round 1 out of 8
Now, we would like you to guess what **the smallest value of X** will be in your group this round and to tell us how confident you are about your guess.

Please, answer these two questions:
What is **your best guess** of the smallest value of X? [*Entry field*]
How confident are you in your guess above?
Not at all confident o o o o o o o Very confident
Press the red button below to confirm your responses.
Confirm [*Button*]

EARNINGS TABLE

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

****

Fig. SM4.3 below reproduces the input screen for the first period belief and confidence elicitation.

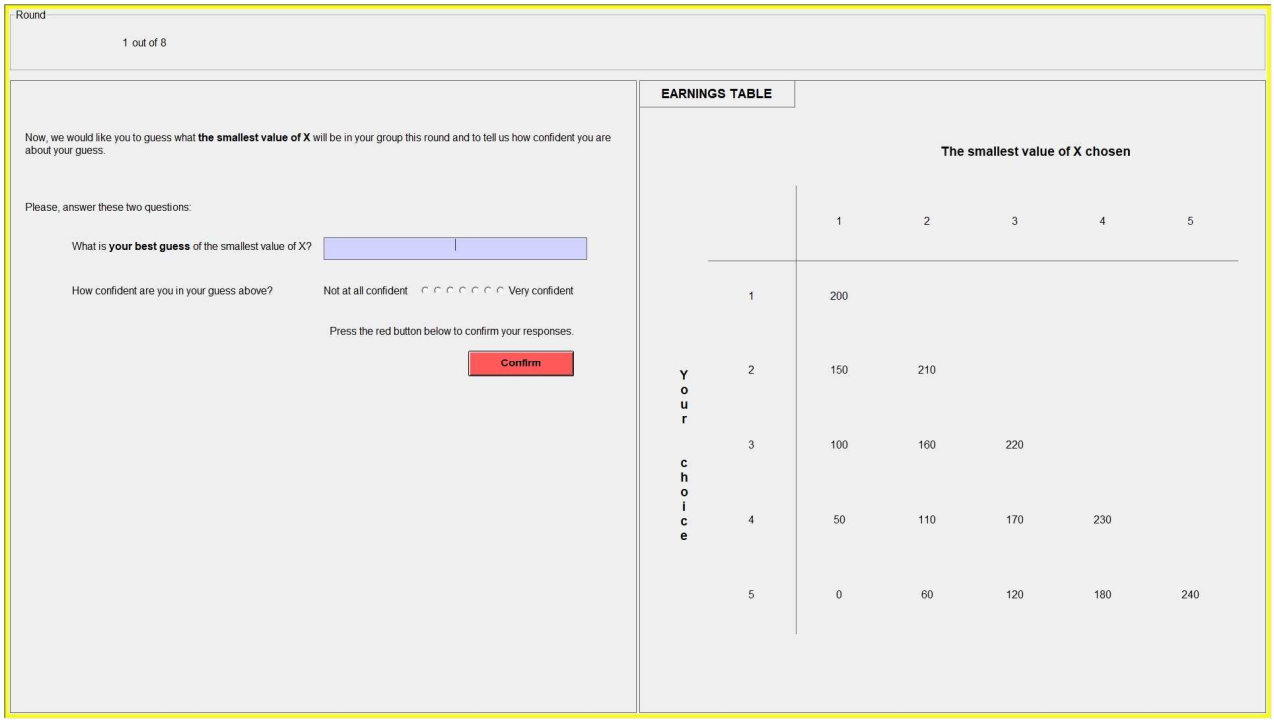


FIG. SM4.3.—Belief elicitation in period 1: input screen.

In each period, after each participant had entered their guess of the smallest value of X and stated their confidence, 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

	×		×	Aisle	×	**Person X**	×	**YOU**

[Note: The '×' above indicated a computer switched-off and thus not in use]

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

	×		×	Aisle	×	**Person X**	×	**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**.

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

Confirm [*Button*]

****

Fig. SM4.4 below reproduces the screenshot eliciting the IOS rating.

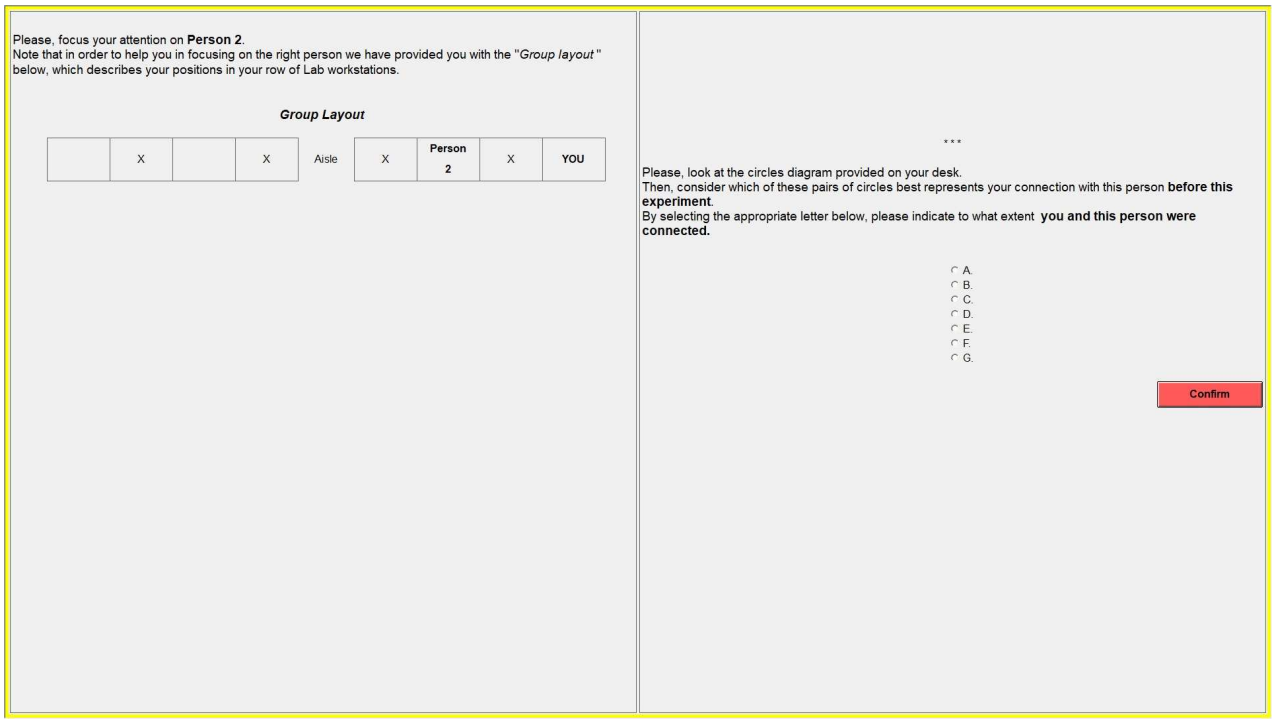


FIG. SM4.4.—IOS scale: input screen. Rater: Person 1; Focus Person: Person 2. The group was allocated to four workstations in a row having eight workstations of which only four were switched on. The positions of the switched-off workstations were marked by '×'.

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

	×		×	Aisle	×	**Person X**	×	**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

	×		×	Aisle	×	**Person X**	×	**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
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

1 2 3 4 5 6 7

15. Tender	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

	×		×	Aisle	×	**Person X**	×	**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
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
20. Perturbed	o	o	o	o	o	o	o
21. Heavy-hearted	o	o	o	o	o	o	o
22. Sorrowful	o	o	o	o	o	o	o

1 2 3 4 5 6 7

23. Bothered	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
24. Kind	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
25. Sad	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
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]

****

Task 6 is an adaptation of the z-Tree implementation by Crosetto et al. (2012) of the paper-based Social Value Orientation (SVO) Slider Measure by Murphy et al. (2011).

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

****

Instructions for Task 6

In this task we do not speak in terms of pounds but of points. Your earning in points will be converted into pounds at the following rate 1 point = 2 pence.
The task involves a set of 15 decision situations. You have been randomly paired with another person from your group. You will be making a series of decisions about allocating resources (in points) between you and this other person in each of the 15 decision situations. Here is an example.

In this example, if you pick the button on the far left, you receive 30 points and the other person receives 80 points. If you pick the button furthest to the right, you get 70 points and the other person gets nothing.
In all 15 decision situations, your task is to indicate the distribution you most prefer by clicking the relevant radio button. The example illustrates the case where you chose to receive 50 and for the other person to receive 40.
You are making decisions for you and Person X. Another person (not Person X) in your group will be making decisions for themselves and you in this task.
All of your decisions are completely confidential.
Please think carefully about each decision. One decision situation will be randomly selected at the end and paid out to you and the other person according to your decision in that situation.
There are no right or wrong answers, this is all about personal preferences.
After you have made your decision, click on the respective button and the resulting distribution of points will be shown on the right. You can revise your choice until you click on the OK button. As you can see, your choices will influence both, the amount of money you receive as well as the amount of money the other person receives.
Please click the button to go to the decision screen.
Continue to the decision screen [*Button*]

****

Following Task 6 instructions, subjects faced the Social Value Orientation (SVO) elicitation task, which entailed 15 questions presented as question 1 reported below:

****

There will be 15 questions.
For each question, please indicate the distribution you most prefer.

1 of 15

You receive	100	94	88	81	75	69	63	56	50
	o	o	o	o	o	o	o	o	o
Person X receives	70	74	78	81	85	89	93	96	100

You receive 0

Person X receives 0

OK [*Button*]

****

Fig. SM4.5 below reproduces the screenshot eliciting participants' risk attitudes.

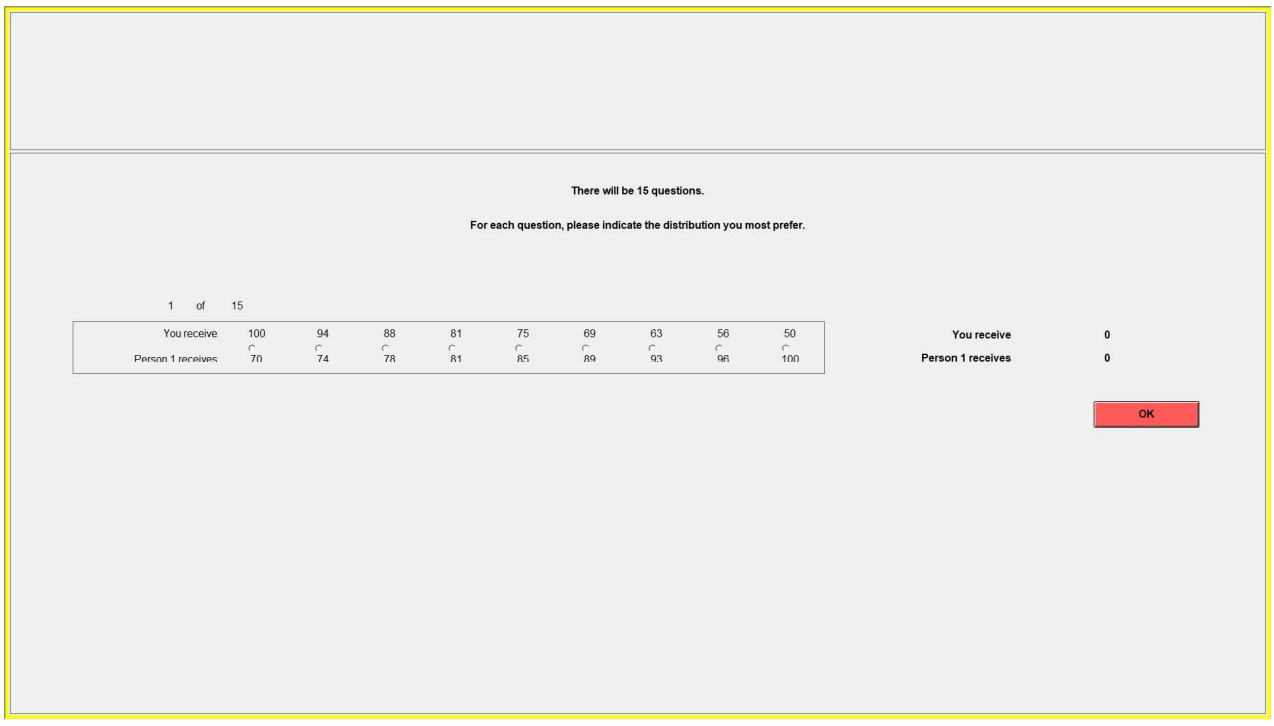


FIG. SM4.5.—SVO elicitation: question 1 out of 15 – input screen.

After participants made their 15 choices, their computer screen would have reported their feedback as described below:

****

The following choice of yours was selected for payment: Choice [*Output field*]

You assigned to yourself: [*Output field*]
You assigned to Person X: [*Output field*]
The following choice of some other person was selected for payment: Choice [*Output field*]
Another person assigned to you: [*Output field*]
Another person assigned to self: [*Output field*]
In total, you receive: [*Output field*]

OK [*Button*]
****

Following Task 6 feedback, they were prompted to start the control tasks reported in the next subsection.

SM4.d. The Control Tasks

The control tasks took the form of a post-experimental, computerized, survey. This is described in this section as it was implemented in Study 2. (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?
[Dichotomous choice: *Male; Female*]

[*2.*] How old are you?

[*3.*] Nationality
[Multiple choice: *British; EU; Other*]

[*4.*] How many siblings do you have?

[*5.*] If you are a student, what is your subject?
[Multiple choice: *Arts and Education; Business economics; Economics; Engineering; Law; Medicine and Health Sciences; Other Social sciences; Politics and International Relations; Science*]

[*6.*] When you were 16 years of age, what was the income of your parents in comparison to other families in your country?
[Multiple choice: *Far below average; Below average; Average; Above average; Far above average*]

[*7.*] How large was the community where you have lived the most time of your life?
[Multiple choice: *Up to 2'000 inhabitants; 2'000 to 10'000 inhabitants; 10'000 to 100'000 inhabitants; More than 100'000 inhabitants*]

[*8.*] How many people live in your household (please include yourself)?

[*9.*] How large is your monthly budget (expenses for accommodation already detracted)?

[*10.*] What share of your monthly expenses you finance yourself?

[*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
[Multiple choice: *No membership; Member; Active member; On the board*]

Music group
[Multiple choice: *No membership; Member; Active member; On the board*]

Political party
[Multiple choice: *No membership; Member; Active member; On the board*]

Lobby group (e.g. student association)
[Multiple choice: *No membership; Member; Active member; On the board*]

Non-profit institution
[Multiple choice: *No membership; Member; Active member; On the board*]

Other kind of clubs
[Multiple choice: *No membership; Member; Active member; On the board*]

[*12.*] Are you religious?
[*Likert scale from 1 "Not at all religious" to 7 "Very religious"*]

[*13.*] Please indicate your political attitude in the following scale.
[*Likert scale from 1 "Left" to 7 "Right"*]

[*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"*]

[*15.*] How satisfied you expect to be in five years time?
[*Likert scale from 1 "Not at all satisfied" to 10 "Absolutely satisfied"*]

[*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"*]

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"*]

[*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"*]

[*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"*]

[*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"*]

[*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"*]

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

[*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?
[Dichotomous choice: *"Would take advantage of you"; "Would try to be fair"*]

[*24.*] Would you say that most of the time people try to be helpful, or that they are mostly just looking out for themselves?
[Dichotomous choice: *"Try to be helpful"; "Just look out for themselves"*]

[*25.*] You can't count on strangers anymore.
[Dichotomous choice: *"More or less agree"; "More or less disagree"*]

[*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"*]

[*28.*] How many other participants (not including your friends) of the experiment do you know by name?

[*29.*] Are you going to share your earnings from the experiment with (any of the) other members of your group, after the experiment?
[*Likert scale from 1 "Certainly not" to 7 "Certainly"* – This has a slightly modified wording relatively to the respective Study-1 question]

[*30.*] Did the possibility that you might share your earnings with other members of your group, after the experiment, play a role in the decision you took?
[*Likert scale from 1 "No role at all" to 7 "Decisive role" "* – This has a slightly modified wording relatively to the respective Study-1 question]

[*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"*]

Your show-up fee for coming in time is 2.50 (in pounds).
Whereas, your experimental earnings are [*Output field*] (in pounds).
Hence, your total earnings in the experiment is [*Output field*] (in pounds), which is the amount you will receive as payment.
Please, fill in the receipt form and sing it off, before pressing the button "Continue".
Thank you very much for your participation!
(Please, kindly remain seated and in silence until called forward for payment by your participation number)

SM5: Two-week Study: Additional Graphical Analysis

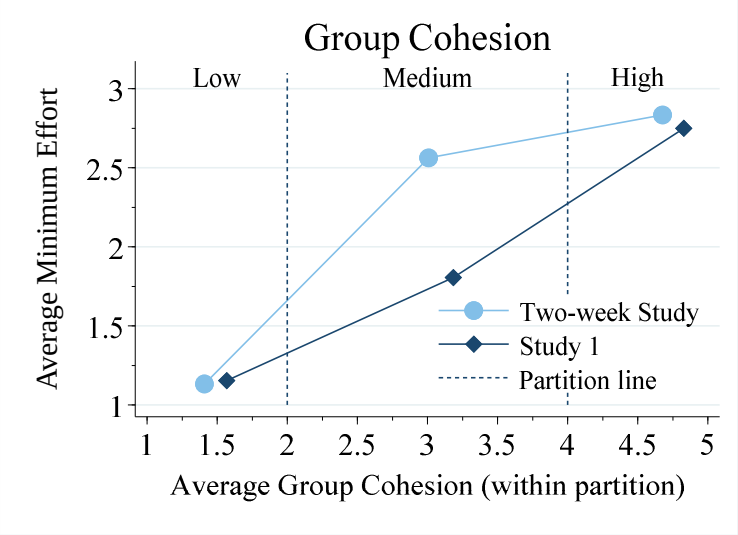


FIG. SM5.1. — Minimum Effort and Group Cohesion in Two-week Experiment and Study 1 (Group-level data from Period 8). For comparability, the average group cohesion for the Two-week Study is based on week 2 group cohesion. "Low Cohesion" partition: *group cohesion* ∈ [1, 2]; "Medium Cohesion" partition: *group cohesion* ∈ (2, 4]; "High Cohesion" partition: *group cohesion* ∈ (4, 7]. The behavior of medium cohesion groups is, seemingly, affected by prior elicitation of oneness. We conjecture the following explanation. 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 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.

SM6: Study 1 and 2: Additional Econometric Analysis

TABLE SM6.1

TEAM PERCEPTION, TRUST AND BEHAVIORAL RISK ATTITUDE AS POSSIBLE ALTERNATIVES TO GROUP COHESION: ORDERED PROBIT REGRESSIONS

Panel A – Team perception: Study 1 and Study 2 combined				
Dep. variable: Minimum Effort	(1)	(2)	(3)	(4)
Team perception	0.267***	0.204***	0.223***	0.183***
	(0.056)	(0.061)	(0.069)	(0.071)
Homophily index			2.040	1.103
			(1.819)	(1.784)
Share (principal component)		0.235***		0.226***
		(0.081)		(0.082)
Study 2 (dummy variable)	-0.274	-0.376*	-0.306	-0.389**
	(0.191)	(0.194)	(0.193)	(0.196)
Log-likelihood	-1264.9	-1234.6	-1260.1	-1233.3
# level 1 (resp. 2) units	1072 (134)	1072 (134)	1072 (134)	1072 (134)
Panel B – Trust: Study 1 and 2 combined				
Dep. variable: Minimum Effort	(5)	(6)	(7)	(8)
Trust (principal component)	0.035	-0.001	-0.013	-0.025
	(0.074)	(0.078)	(0.071)	(0.075)
Homophily index			5.067***	3.315**
			(1.484)	(1.573)
Share (principal component)		0.327***		0.270***
		(0.073)		(0.079)
Study 2 (dummy variable)	-0.223	-0.360*	-0.297	-0.386**
	(0.188)	(0.189)	(0.195)	(0.194)
Log-likelihood	-1338.1	-1270.7	-1295.6	-1255.3
# level 1 (resp. 2) units	1072 (134)	1072 (134)	1072 (134)	1072 (134)
Panel C – Behavioral Risk Attitude: Study 1 and 2 combined				
Dep. variable: Minimum Effort	(9)	(10)	(11)	(12)
Behavioral Risk Attitude	0.103*	0.091	0.103	0.093
	(0.060)	(0.061)	(0.065)	(0.064)
Homophily index			5.042***	3.281***
			(1.468)	(1.537)
Share (principal component)		0.320***		0.261***
		(0.075)		(0.080)
Study 2 (dummy variable)	-0.281	-0.427*	-0.385**	-0.470**
	(0.190)	(0.192)	(0.187)	(0.193)
Log-likelihood	-1328.0	-1262.7	-1285.2	-1247.5
# level 1 (resp. 2) units	1072 (134)	1072 (134)	1072 (134)	1072 (134)

NOTE.—Data from Periods 1 to 8. Explanatory variables are at group level. Variable definition and construction are in SM3 above. Period dummies (always included, relative to Period 1) are significantly negative (at $p < 0.05$). Controls for individual effects: nested random effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

TABLE SM6.2

GLLAMM ORDERED PROBIT REGRESSIONS OF MINIMUM EFFORT ON GROUP COHESION AND CONTROL VARIABLES

Panel A – Study 1						
Dep. variable: Minimum Effort	(1)	(2)	(3)	(4)	(5)	(6)
Group cohesion	1.225***	0.972***			1.360***	1.204***
	(0.316)	(0.325)			(0.396)	(0.392)
Homophily index			11.012*	4.503	-3.966	-7.467
			(5.714)	(5.929)	(6.933)	(6.992)
Share (principal component)		0.671**		0.947***		0.738**
		(0.305)		(0.321)		(0.314)
Log-likelihood	-376.8	-374.8	-381.4	-378.2	-376.6	-374.3
# level 1 (resp. 2) units	520 (65)	520 (65)	520 (65)	520 (65)	520 (65)	520 (65)
Panel B – Study 2: Pre-registered replication independently conducted at the University of Birmingham						
Dep. variable: Minimum Effort	(7)	(8)	(9)	(10)	(11)	(12)
Group cohesion	1.035***	0.763***			0.919***	0.681**
	(0.243)	(0.263)			(0.300)	(0.312)
Homophily index			16.499***	10.369*	4.198	3.170
			(5.190)	(5.679)	(6.298)	(6.514)
Share (principal component)		0.609**		0.865***		0.598*
		(0.302)		(0.293)		(0.307)
Log-likelihood	-345.7	-344.0	-349.5	-345.9	-345.5	-343.9
# level 1 (resp. 2) units	552 (69)	552 (69)	552 (69)	552 (69)	552 (69)	552 (69)
Panel C – Study 1 and 2 combined						
Dep. variable: Minimum Effort	(13)	(14)	(15)	(16)	(17)	(18)
Group cohesion	1.097***	0.826***			1.089***	0.873***
	(0.189)	(0.201)			(0.234)	(0.237)
Homophily index			13.477***	7.059*	0.198	-1.837
			(3.685)	(3.976)	(4.476)	(4.549)
Share (principal component)		0.629***		0.895***		0.643***
		(0.210)		(0.211)		(0.214)
Study 2 (dummy variable)	-0.837*	-1.015**	-0.818	-1.053**	-0.841*	-1.000**
	(0.499)	(0.504)	(0.507)	(0.512)	(0.500)	(0.505)
	All <1%					
Log-likelihood	-741.9	-738.2	-750.8	-743.6	-741.9	-738.1
# level 1 (resp. 2) units	1072 (134)	1072 (134)	1072 (134)	1072 (134)	1072 (134)	1072 (134)

NOTE.—Data from Periods 1 to 8. Explanatory variables are at group level. Variable definition and construction are in SM3 Above. Period dummies (always included, relative to Period 1) are significantly negative (at $p < 0.01$). Controls for individual effects: nested random effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

TABLE SM6.3

ORDERED PROBIT REGRESSIONS OF MINIMUM EFFORT ON AVERAGE ONENESS AND CONTROL VARIABLES

Panel A – Study 1				
Dep. variable: Minimum Effort	(1)	(2)	(3)	(4)
Group average oneness	0.358***	0.298***	0.385***	0.358**
	(0.096)	(0.111)	(0.135)	(0.140)
Homophily index			-0.752	-1.865
			(2.590)	(2.557)
Share (principal component)		0.186		0.204
		(0.128)		(0.128)
Log-likelihood	-657.3	-647.4	-657.0	-645.8
# level 1 (resp. 2) units	520 (65)	520 (65)	520 (65)	520 (65)
Panel B – Study 2: Pre-registered replication independently conducted at the University of Birmingham				
Dep. variable: Minimum Effort	(5)	(6)	(7)	(8)
Group average oneness	0.427***	0.344***	0.375***	0.304**
	(0.105)	(0.118)	(0.135)	(0.143)
Homophily index			2.290	1.927
			(3.044)	(2.944)
Share (principal component)		0.205*		0.197*
		(0.118)		(0.115)
Log-likelihood	-561.3	-551.5	-558.6	-549.7
# level 1 (resp. 2) units	552 (69)	552 (69)	552 (69)	552 (69)
Panel C – Study 1 and 2 combined				
Dep. variable: Minimum Effort	(9)	(10)	(11)	(12)
Group average oneness	0.395***	0.319***	0.376***	0.320***
	(0.069)	(0.079)	(0.095)	(0.100)
Homophily index			0.622	-0.048
			(2.010)	(1.941)
Share (principal component)		0.199**		0.200**
		(0.085)		(0.084)
Study 2 (dummy variable)	-0.319*	-0.401**	-0.327*	-0.400**
	(0.193)	(0.195)	(0.192)	(0.196)
Log-likelihood	-1236.8	-1215.9	-1236.4	-1215.9
# level 1 (resp. 2) units	1072 (134)	1072 (134)	1072 (134)	1072

NOTE.—Data from Periods 1 to 8. Explanatory variables are at group level. Variable definition and construction are in SM3 above. Period dummies (always included, relative to Period 1) are significantly negative (at $p < 0.05$). Controls for individual effects: nested random effects. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

TABLE SM6.4

MEDIATION ANALYSIS: ESTIMATING SIMULTANEOUS LINEAR EQUATIONS

Equation 1	
Dep. variable: Minimum Effort	(1)
Beliefs	0.865***
	(0.025)
Social Preferences	-0.002
	(0.039)
Group Cohesion	0.072**
	(0.033)
Equation 2	
Dep. variable: Beliefs	(2)
Social Preferences	0.254***
	(0.080)
Group Cohesion	0.280***
	(0.100)
Equation 3	
Dep. variable: Social Preferences	(3)
Group Cohesion	0.532***
	(0.068)
Log-likelihood	-4272.3
# level 1 (resp. 2) units	552 (69)

NOTE.—Study-2 data from Periods 1 to 8. All coefficients are standardized. Clustering at group level applied. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

TABLE SM6.5

STUDY 1 AND 2 COMBINED: ORDERED PROBIT REGRESSION OF GROUP COHESION ON PRIOR INTERACTIONS AND EMPATHY – A FURTHER TEST OF GROUP COHESION'S CONSTRUCT VALIDITY

Dependent variable:	Group Cohesion	
Prior group interactions	0.632***	(0.074)
Empathy index	0.673***	(0.123)
Log-likelihood	-403.7	
# Observations	134	

NOTE.—Standard errors in parentheses. *** $p < 0.01$.

SM References

Crosetto, Paolo and Weisel, Ori, and Winter, Fabian. 2012. "A Flexible z-Tree Implementation of the Social Value Orientation Slider Measure (Murphy et al. 2011) – Manual." Jena Economic Research Paper # 2012-062.

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.

Murphy, Ryan O., Ackermann, Kurt A., and Handgraaf, Michel J.J. 2011. "Measuring Social Value Orientation." *Judgment and Decision Making*, 6 (8): 771-781.

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.