Measuring "Group Cohesion" to Reveal the Power of Social Relationships in Team Production†

Simon Gächter[1,2,3], Chris Starmer[1,*] and Fabio Tufano[1,*].

[1] CeDEx, School of Economics, University of Nottingham, NG7 2RD
[2] IZA, Bonn, Germany
3 CESifo, Munich, Germany
* Corresponding authors (chris.starmer@nottingham.ac.uk, fabio.tufano@nottingham.ac.uk).

25 July 2022

We introduce "group cohesion" to study the economic relevance of social relationships in team production. We operationalize measurement of group cohesion, adapting the "oneness scale" from psychology. A series of experiments, including a pre-registered replication, reveals strong positive associations between group cohesion and performance assessed in weak-link coordination games, with high-cohesion groups being very likely to achieve superior equilibria. In exploratory analysis, we identify beliefs rather than social preferences as the primary mechanism through which factors proxied by group cohesion influence group performance. Our evidence provides proof-of-concept for group cohesion as a useful tool for economic research and practice. JEL Classification: C92, D91

Acknowledgments. We are grateful to Robert Cialdini and John K. Maner for sharing 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 with experiments. We have benefited greatly from discussions with: 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 thank participants at various conferences, workshops, and seminars for their valuable feedback. This work was supported by the Economic and Social Research Council [grant numbers ES/K002201/1, ES/P008976/1], the European Research Council [grant numbers ERC-AdG 295707 COOPERATION, ERC-AdG 101020453 PRINCIPLES], and the Italian Ministry of Education. The research was approved by the School of Economics Research Ethics Committee (University of Nottingham). The authors declare no relevant conflict of interest.

† This paper reports an extensive programme of work and, at various stages of its development, elements of it were reported in working papers under the following titles: (i.) "The surprising capacity of the company you keep: revealing group cohesion as a powerful factor of team production"; (ii.) "Revealing the economic consequences of group cohesion"; (iii.) "Measuring the impact of social relationships: the value of 'oneness'"; (iv.) "The power of social relations for coordination: the magic of 'oneness'".

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. In this paper, we report an extensive program of conceptual and experimental research building from the arguably plausible idea that the 'qualities' of social relationships within households, firms, or other organizations, collectively constitute an important factor of production. While various strands of existing literature hypothesize that social relationships may matter for economic outcomes (see Section II below), our motivation stems from the absence of any systematic approach to *measuring* the productive value of social relationships. Our primary contribution is to develop and test a measurement tool, based on a new concept of *group cohesion*, with a view to providing foundations for the study of social relationships as factors of production.

We proceed via two main steps. The first is to introduce a simple, but conceptually well-grounded, approach to characterizing *any* real group in terms of a group cohesion index, intended as a summary statistic for aspects of social relationships that matter for team production. Our second main step is to test the predictive power of the group cohesion index in a large-scale program of experiments and accompanying analysis investigating (weak-link) team production in real groups that vary in terms of their pre-existing social relationships.

Since group cohesion is a novel concept in economics, in Section III, we explain the rationale for the concept, our approach to measuring it and some of its key properties. To preview briefly, 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" (or sometimes just "cohesion" for brevity) to refer to the state of the aggregate closeness ties within a group. Because 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 uses simple, and very portable methods to assess how close one person feels to another, based on their own self-reports. From a measurement viewpoint,

our innovation is to develop an aggregate statistic, based on oneness, to characterize the set of relationships within a group. Specifically, in our experiments, we ask each group member, privately, to indicate their oneness with every other group member. The group cohesion index is then calculated as an average of individual oneness ratings (full details of our measurement approach and its psychological foundations are in Section III.A).

This seemingly modest measurement innovation generates a tool with considerable predictive power: Across a set of six experiments (see Appendix Table 3 for a summary of key aspects), we demonstrate that group cohesion is strongly associated with group outcomes. We explain the main experimental setup in Section IV. A key feature is that we study the behavior of *real* groups – not artificially created ones – achieved by recruiting groups of friends to participate. Hence, we observe real closeness ties based on pre-existing sociological and psychological characteristics that are absent (by construction) in groups set up on the spot in the lab, including those using minimal group manipulations (Goette, Huffman and Meier (2012)).[‡] As we will show, measured cohesion tracks tangible sociological features of the real groups we study (see Section III.B).

Our workhorse to study group outcomes is a weak-link game, chosen to model coordination problems endemic to real organizations and teams (e.g., Camerer and Weber (2013)). In our version of the weak-link game, 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 reflecting two dimensions of group success: *coordination* (matching the effort level of other group members) and *cooperation* (achieving Pareto-superior equilibria). Building on the approach of Brandts and Cooper (2006), we designed our game to be "harsh" in the sense that groups lacking pre-existing social relationships would be expected to collapse to the Pareto-worst equilibrium.

Section V presents the key behavioral patterns in our data. We identify a strong positive association between group cohesion and performance and, while the likelihood of coordinating on some equilibrium is largely independent of the level of cohesion, it is crucial for equilibrium selection: low cohesion groups usually descend rapidly to the worst Pareto-ranked equilibrium; high cohesion groups typically achieve higher Pareto-ranked equilibria. We replicate these patterns via an independently conducted, pre-registered, experiment (Study 2, Table 3,

[‡] In the taxonomy of Charness et al. (2013), our experiments would classify as "extra-laboratory experiments". However, a more apt label could be "field-in-the-lab experiment" because we bring naturally occurring groups of friends into a laboratory setting.

Appendix). We also confirm that our results are robust to the timing of oneness measurement (before or after play of the weak-link game), eliminating the interpretation that experience of game play explains variation in cohesion.

While our results clearly establish that groups with high cohesion systematically outperform low cohesion groups, one must be cautious about causal interpretation. Ultimately, our results are correlational, but we think a plausible interpretation of our findings is that group cohesion is a summary statistic for tangible features of real groups that matter, causally, for team production (at least in the context of the weak-link settings we study). Interpreted in that way, the group cohesion index as a new tool would be much less valuable if one could predict group performance, just as well, using a small number of easily measured group characteristics; our results, however, cast doubts on the prospects for doing that.

In Section VI, we present econometric analysis showing that group cohesion is a powerful and dominant predictor of group performance even when controlling for a large range of measured group characteristics – moreover, those characteristics become insignificant as predictors of group outcomes, once cohesion is present as a regressor. In the last game period, the model predicted effects of cohesion on group outcomes are also substantial: minimally cohesive groups are almost certain to collapse to a minimum effort of 1; maximally cohesive groups almost never fall to a minimum effort of 1; and large financial incentives are needed to promote the levels of effort expected for high cohesion groups.

In Section VII, we discuss the interpretation of our results considering two main avenues. First, we consider the possibility that, because our experiments involve groups of friends, the association between effort and cohesion might be explained by subjects having planned to share their earnings with participating friends. We test and discount this as a plausible account of our main findings. Second, exploiting data on participants' beliefs and social preferences gathered in our replication study, we explore the extent to which the association between group cohesion and minimum effort is mediated through beliefs or social preferences. In contrast to results found elsewhere (e.g., Chen and Chen (2011)), we find that the effects of cohesion operate mainly via the channel of beliefs with only a limited influence of social preferences.

To preview our conclusion of Section VIII, our studies establish proof of concept for group cohesion as a useful new tool of economic analysis to capture and reveal the, previously hidden, power of social relationships as factors of production.

II. Related Literature and Our Contribution

Before presenting the substance of our paper we briefly place it in the literature. In the broadest sense, we contribute to the literature on social capital (e.g., Putnam (2000); Glaeser, Laibson and Sacerdote (2002)) by tackling one of its central problems. In a recent typology of that literature, Jackson (2020) argues that "[m]easuring various forms of social capital is especially difficult as they are dependent upon relationships between people, which are often intangible and only indirectly observed" (p. 333). We demonstrate how (social) relationships can be observed and measured to provide quantitative assessments of the (psychological) quality of social network links (Goyal (2005)), thereby providing a micro-foundation of social capital. We do this by introducing the novel psychological concept of group cohesion. As we will explain, group cohesion builds on the concepts of "relationship closeness" and "oneness". These concepts are firmly established in the psychology literature (see Section III.A) but are less considered in economics, with the possible exception of "social distance" (e.g., Akerlof (1997)). In the experimental economics literature, social distance has mainly been 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)). By contrast, we *measure* the closeness of relationships between group members and construct the concept of group cohesion on such measurements. To our knowledge, this is an entirely new approach in economics.

Another contribution is to the experimental literature on coordination games, which – following seminal papers by van Huyck, Battalio and Beil (1990; 1991) – has largely studied coordination among anonymous individuals without considering the role of social relationships. This research (see Cooper and Weber (2020) for a recent survey), highlights primarily the importance of structural features that facilitate coordination on efficient equilibria such as: communication (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 socio-psychological property of group cohesion is an independent and powerful predictor of group outcomes.

We also contribute to a growing literature on the economic impact of groups and group identity (Charness and Chen (2020)). This includes studies investigating in-group favoritism (e.g., Currarini and Mengel (2016)); interactions among friends (e.g., Glaeser et al. (2000); Leider et al. (2009); Babcock et al. (2019); Chierchia, Tufano and Coricelli (2020); Gächter et

al. (2022)); and the role of identity in organizations (e.g., Akerlof and Kranton (2005); Ashraf and Bandiera (2018)), including social-psychological dimensions of employment relationships (e.g., Baron and Kreps (2013)). Our work builds most directly on prior experimental work which has established the impact of group identity on behavior in various contexts including in prisoner's dilemma and battle of the sexes games (Charness, Rigotti and Rustichini (2007)), in trust games (Hargreaves Heap and Zizzo (2009)) and in weak-link games (Chen and Chen (2011)). While the last of these comes closest to our work in studying weak-link games, relative to all three studies, our work breaks new ground: we study *real groups*, not artificially constructed ones; and we do this for the novel purpose of developing and providing proof of concept for a tool to measure the quality of behaviorally relevant features of extant socio-psychological relationships within real groups.

III. Group Cohesion in Real Groups

Since group cohesion is a novel concept in economics, we devote subsection A to explaining the concept, its roots in established psychological literature and our approach to its measurement. Subsection B shows that measured group cohesion passes a basic test of construct validity in varying coherently with tangible sociological properties of real groups.

A. Group Cohesion: psychological foundations and measurement

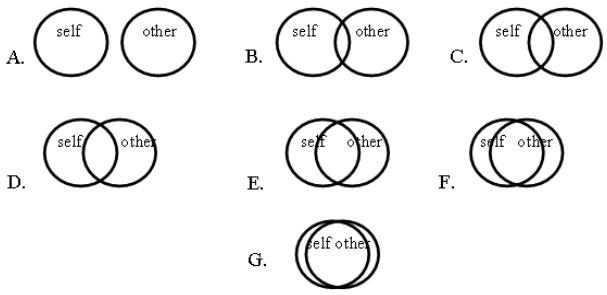
Our study involves the development and application of a new tool: A simple and portable measure of group cohesion designed to summarize the social and psychological relationships that exist between members of any group. To this end, we build on an established literature which has developed tools to measure the nature and strength of *bilateral* relationships between pairs of individuals. This literature demonstrates that important 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 and extends this literature by assuming that important aspects of relationships that exist within groups can be summarized in terms of the *set of pairwise closeness relationships within them*. 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 foundation, we now describe the key properties of tools for measuring bilateral 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 assess, 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, et al. (1992) proposed a simple tool: the "Inclusion of the Other in the Self" (IOS) scale depicted in Fig. 1a. The IOS scale "is hypothesized to tap people's *sense* of being interconnected with another. That sense may arise from all sorts of processes, conscious or unconscious" (Aron, et al. (1992), p. 598). Essentially, it aims to measure relationship closeness without examining its detailed determinants (i.e., frequency or diversity of activities; strength of mutual influence).

Aron, et al. provide statistical evidence that the IOS scale successfully tracks key dimensions of relationship closeness: people tend to pick more overlapping pairs of circles for a given other, the more frequent or diverse their interactions, and the stronger their perceptions of mutual influence. 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. Their measure also correlates strongly with the IOS scale. Together, these results make the IOS scale a very promising tool for our purposes. It also has the decisive advantage of being intuitive for respondents and very simple to implement.

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

A. ☐ B. ☐ C. ☐ D. ☐ E. ☐ F. ☐ G. ☐



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

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

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

b. The We Scale

Figure 1. Oneness elicitation as explained to the participants.

Since our research relies critically on the IOS scale, in a background paper (Gächter, et al. (2015)), via a study with 772 subjects, we assessed the psychometric validity of the IOS scale for a wide range of relationships (from strangers to close friends), by testing whether we could replicate key findings in the foundational psychological research that validated the IOS scale as a reliable predictor of relationship closeness. Our results replicate, remarkably closely, the correlations of the IOS scale with the Relationship Closeness Index reported by Aron et al. The IOS scale also varies coherently with the form of the relationship (lowest for acquaintances, medium for friends, and highest for close friends), with the Personal Acquaintance Measure of Starzyk, et al. and with Rubin's Loving and Liking Scale. In Gächter, et al. (2015), we also find that the principal components of the questionnaire-based measures correlate strongly (0.85) with the IOS scale. Hence, we conclude that the IOS scale is a psychologically meaningful and reliable tool for measuring bilateral relationship closeness.

In our measurement of relationship closeness, we follow Cialdini, et al. (1997) who combine the IOS scale with the "We scale," depicted in Fig. 1b. The Cialdini et al. measure is calculated as the average of responses on these two scales. They call this the "oneness scale," which they interpret as reflecting a "sense of shared, merged, or interconnected personal identities" (p. 483). In Gächter, et al. (2015), we confirmed Cialdini et al.'s claim that oneness correlates slightly better with the questionnaire-based measures than the IOS scale alone and, hence, we use the oneness scale for our analysis.

We deployed the oneness scale as follows (wider procedural details are in Section IV). Subjects participated as groups of four and each person rated three other visually identified group members, separately and privately, on the IOS and We scales as depicted in Fig. 1; group members knew they would not receive feedback about each other's ratings. Both IOS and We scale responses are scored on a scale from 1 to 7. Oneness is the average of the two measures and hence ranges from 1, "lowest oneness", to 7, "highest oneness."

Since groups contain four people, who each rate the three other members in their group, any group generates twelve bilateral oneness ratings. We construct our group cohesion index by selecting, for each group member, the oneness score for the person they rated lowest. We then compute group cohesion as the average of these four scores. Hence, our index can be thought of as summarizing the minimum envelope of oneness in a group. Our results are not sensitive, however, to different ways of averaging the individual oneness reports (see Section VI.B).

B. Group Cohesion and the Sociological Properties of Naturally Occurring Groups

In much of this paper, we focus on whether or how well group cohesion predicts performance in stylized "production tasks". Before pursuing this, however, we briefly probe the validity of our measurement tool via some simple tests examining whether measured cohesion varies as expected with tangible, socio-demographic, features of the groups in our experiments.

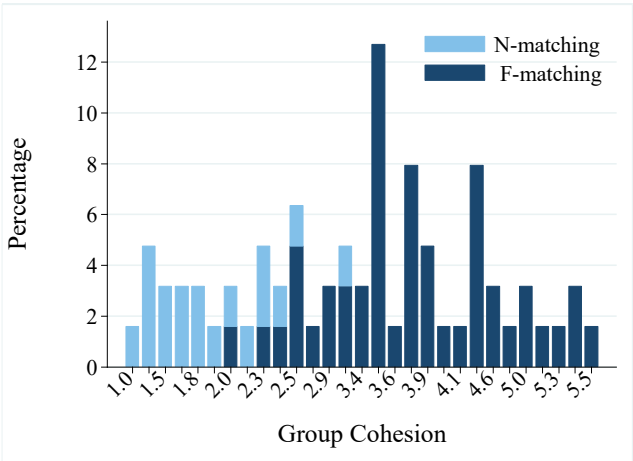


Figure 2. The distribution of group cohesion under F- and N-matching. The N-matching bars are stacked over the F-matching ones.

The simplest approach to this exploits the procedures we used to assemble groups. To create variation in how well members of groups knew one another prior to our experiments, we recruited participants as groups of four self-selecting 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) (see Section IV for further details). If cohesion is tracking the pre-existing relationships within groups, we should expect that already existing groups (F-matching) will tend to have higher measured group cohesion than the ones we constructed fresh (N-matching). Fig. 2 plots the distribution of measured group cohesion separately for N- and F-matching. It is evident that group cohesion tends to be higher in the F-matching groups as compared to N-matching groups (means are 3.81 and 1.84, resp.; the

distributions differ according to a Mann-Whitney test, $z = 5.816$; $p < 0.001$). Note that, with measured group cohesion ranging from 1 to 5.5, there is good scope for observing its association with group behavior, if such association exists.

For a more sophisticated test of construct validity, we use individual-level data on 15 characteristics of our participants, collected via post-experimental questionnaires. The characteristics range from self-reports of relatively concrete variables (e.g., age or gender) through to more subjective self-assessments of dispositional traits (e.g., political attitudes or happiness). An established literature related to "homophily" and the sociology of friendship (e.g., Baccara and Yariv (2013); Dunbar (2018); McPherson, Smith-Lovin and Cook (2001)) shows that "like-befriends-like" hence members of self-assorted groups are expected to be more similar in terms of socio-demographic characteristics than members of other groups. This is clearly true of our self-selected groups (i.e., the F-matching groups): Based on both parametric and non-parametric tests, the null of equal variance between and within F-matching groups is rejected in the expected direction at $p < 0.05$ for 11 of the 15 characteristics (see Table SM2.1 in the supplemental materials, henceforth "SM"). In contrast, no significant differences are found for N-matching groups. This analysis demonstrates that homophily is an indicator of pre-exiting relationships among group members. Hence, if group cohesion measures what we intend, we should expect that group cohesion and group homophily will be correlated. To test this prediction, we construct a simple *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. We explain the construction of the Homophily Index in detail in Section VI, where it features as a control variable. For now, however, we note that 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. We take this as reassuring evidence that, as well as being a simple, intuitive, and portable group-level statistic capturing bilateral assessments or relationship closeness, the group cohesion index also reflects homophily within groups, consistent with the literature on the sociology of friendship.

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), it has been widely studied in the lab, partly because it represents a form of coordination problem endemic to organizations (Camerer and

Weber (2013)). A classic example is workers preparing an aircraft for takeoff: the plane can only leave once the *slowest* worker has fulfilled their task (Knez and Simester (2001)).

We use a version of the weak-link game due to Brandts and Cooper (2006). In a group of four, players simultaneously choose one of five "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(·) 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 Payoffs (in points) 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

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 as our baseline setup in the expectation that – in the absence of aids to cooperation (e.g., communication) – low cohesions groups, typical of those used in prior experimental implementations of this game (e.g., Brandts and Cooper (2006; 2007)), would rapidly descend to the worst equilibrium.[§]

[§] Relative to other weak-link settings, this one is "harsh" in the sense defined in SM3.

B. Sampling Strategy and Sequence of Events

Since our goal was to study the performance of real groups, 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, "Study 1") 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 was randomly assigned to become a member of a *different* experimental 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).

Since our setup required participants to both provide oneness ratings of other group members and to play a (repeated) weak-link game, a very important issue is whether the experience of one type of task might affect behavior in the other. We addressed this issue in two ways. First, pilot experiments revealed that measuring oneness before the weak-link game does have some influence on minimum effort. A key question is then whether prior play of the game affects subsequently measured oneness. To test this, we ran a within-subject experiment (172 new subjects; 27 F-matching groups and 16 N-matching ones) conducted in two stages. We refer to this as our "two-week experiment" (see Table 3, Appendix). In week 1, we measured oneness and elicited various individual characteristics. In week 2, the same subjects in the same groups played the weak-link game followed by elicitation of oneness. Since relationship closeness should not change systematically over the course of one week, any systematic 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). 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). This demonstrates an encouraging degree of test-retest reliability at the level of the individual. 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, β = -0.032, z = -0.28, p =

0.783).[**] 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 total accumulated points across the eight periods would be converted to cash at an exchange rate of 500 points = £1.00. For oneness measurements (elicited after game play for reasons explained above) after computerized instructions, each participant was asked to focus on each other group member in turn and to respond, in sequence, to both the IOS scale and the We scale (Fig. 1) tasks. The full experimental instructions are in the supplemental materials (see section SM14).

We recruited participants via ORSEE (Greiner (2015)) and ran the experiments with z-Tree (Fischbacher (2007)) in the CeDEx lab at the University of Nottingham. 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.

[**] 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 game play.

[††] It was essential for our design that subjects knew who their other group members were and, in particular, that subjects in N-groups realised that they were not grouped with their friends. Hence, in verbal instructions we asked them to "pay attention to the composition of their group" (see oral instructions in SM14). This instruction formed a brief part of the overall instructions, given some time in advance of decisions and we did not provide any signal for how subjects should take account of group membership. A reviewer suggested that this instruction might foster an experimenter demand effect. While we cannot definitively reject such a possibility, studies of experimenter effects suggest that their scale is generally modest (e.g., de Quidt et al., 2018). Nevertheless, direct evidence on this point from further research could be useful.

V. Associations between Group Cohesion and Weak-Link Team Production

Before presenting our primary results, we note that our experimental environment is "harsh", as intended, in that groups whose participants have no significant history of prior social interaction tend to quickly gravitate towards the lowest ranked equilibrium of the weak-link game. Using data from the N-matching, we find that, by period 8, 90 percent of groups collapse to minimum effort = 1; only two groups do better, achieving effort levels 2 and 3, respectively.[‡‡] These results confirm existing evidence (e.g., Brandts and Cooper (2006, 2007)) and establish that there is ample scope for improvement in cooperation in our environment, if the factors captured in the group cohesion index measure matter for team production.

A. Group Cohesion, Minimum Effort and Wasted Effort

Fig. 3 presents scatter plots of minimum effort against group cohesion with separate panels for the first and last periods of the weak-link game. Each plot also includes a line of best fit (OLS) and the 95% confidence interval. 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 (see p values in note to Fig. 3).

[‡‡] A long-time horizon does not help low cohesion groups escape cooperation failure. We tested this with 32 fresh participants, recruited individually, who played the game of Table 1 for *50 periods* in 8 fixed groups of four anonymous members (see Appendix). Six groups were trapped in the Pareto-worst equilibrium by period 4; one by period 10; and one by period 22.

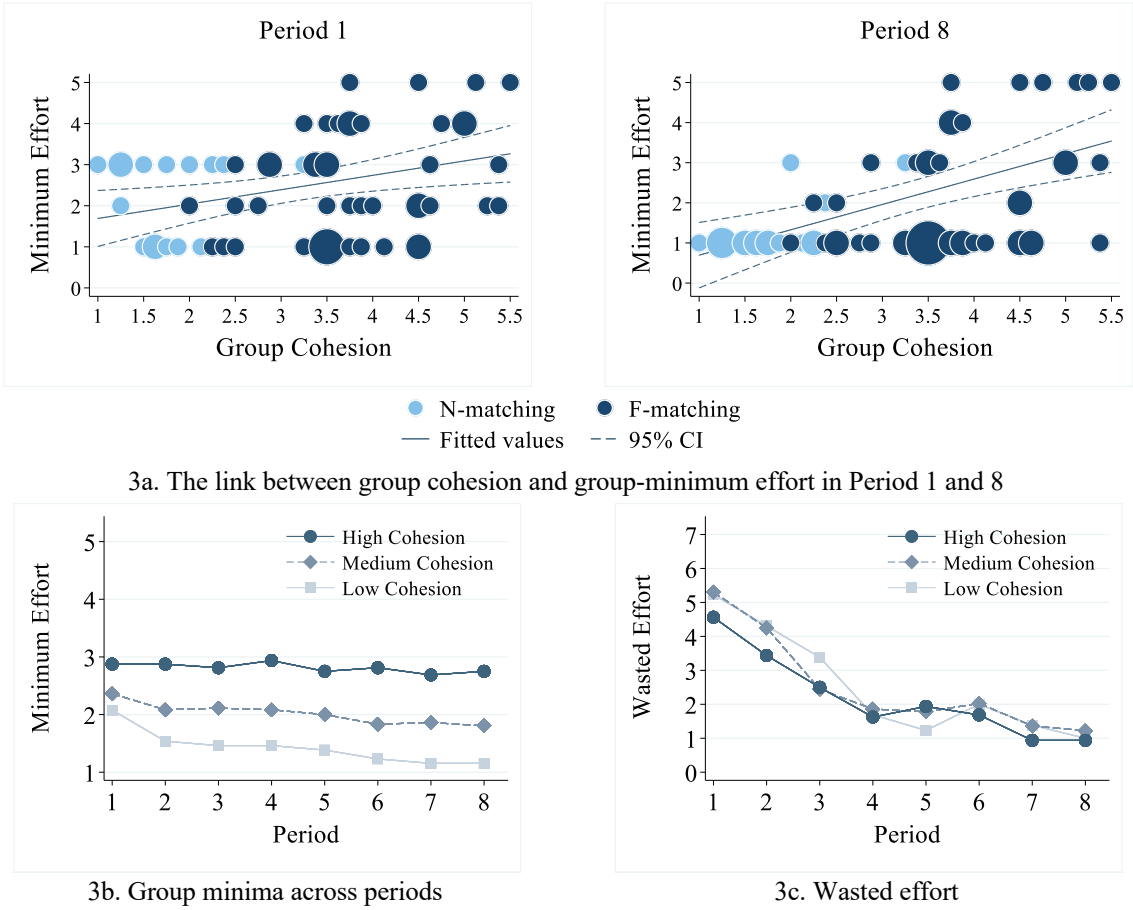


Figure 3. Group Cohesion, Minimum Effort and the Dynamics of Coordination. Fig. 3a: Size of symbols proportional to no. of observations; in Period 1, two N-matching observations are not displayed because they coincide with F-matching circles with coordinates (2.25, 1) and (2.5, 2); in Period 8, one N-matching 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" cohesion 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 and trends in minimum effort 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.

Interestingly, the dynamics of "wasted effort" (i.e., the total of effort in a group above the group minimum in a particular period) seem largely independent of cohesion levels and the uniformly low rates of wasted effort by period 8 imply strong convergence on equilibrium play for all levels of cohesion.[§§] As Fig. 3c shows, average wasted effort in period 1 is around 5 and collapses to about 1 by period 8. The analyses of Figs. 3b and 3c suggest that group cohesion is primarily associated with cooperation (decisions consistent with higher ranked equilibria), with relatively little connection to coordination success (group members coordinating on the same equilibrium, regardless of its ranking).

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 cohesion (i.e., low, medium, and high) introduced in Fig. 3b. These tests show that for both low and high cohesion groups, the achieved levels of minimum effort are statistically indistinguishable across the two experiments. For groups with mid-range cohesion, minimum effort is somewhat higher for the two-week experiment. For both experiments, however, we identify a strong positive association between cohesion and minimum effort, regardless of the timing of the oneness elicitation. This holds regardless of whether we include observations from groups with mid-range cohesion (for more details of analysis see supplemental material, Section SM4).

B. A Pre-Registered Replication

While the results presented in the last subsection are encouraging, they are also novel. Therefore, replicability is of first order importance to establish confidence in the behavioral patterns just reported (e.g., Camerer, Dreber and Johannesson (2019)). We therefore replicated the experiment and report the results in this sub-section. 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[***] for Study 2 and we hired an independent contractor (the University of Birmingham Experimental Economics Laboratory (BEEL)) to implement them. We provided the experimental protocol,

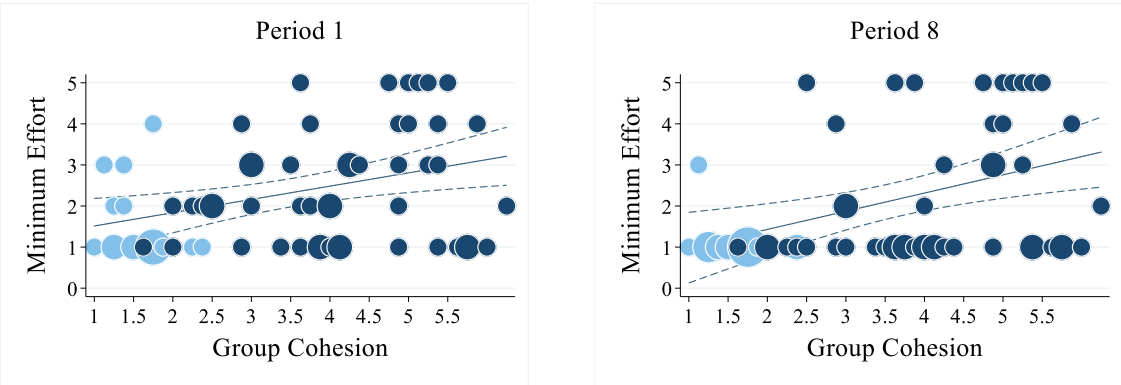
[§§] We find only a weakly significant relationship between (average) group level wasted effort and group cohesion (Spearman's $\rho = -0.227$, $p = 0.069$; $n = 65$).

[***] See https://www.socialscienceregistry.org/trials/3566 (Reg. no. AEARCTR-0003566). Note that we collected one fewer group in the F-matching than planned due to no-shows.

software, and instructions, but we were not involved in data collection. BEEL followed our original recruitment procedures but with a new subject pool from Birmingham University. The protocols and instructions were as for Study 1 except that, to probe the relationship identified in Study 1, we introduced two further sets of measurements. First, subjects' beliefs about the minimum effort in their group were elicited in each round of the weak-link game. Second, the post-experimental questionnaire included incentivized elicitation of "Social Value Orientation" (Murphy and Ackermann (2014)) as a measure of group social preferences. We discuss the details of these measures and the associated results in Section VII.

The main results of Study 2 (276 participants; 49 F-matching groups and 20 N-matching ones) are described in Figure 4. 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 the first and the last period. 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 the dynamics of wasted effort are largely independent of cohesion levels.[‡‡‡] In sum, the results of Study 2 confirm that group cohesion has a replicable association with cooperation in the weak-link game.

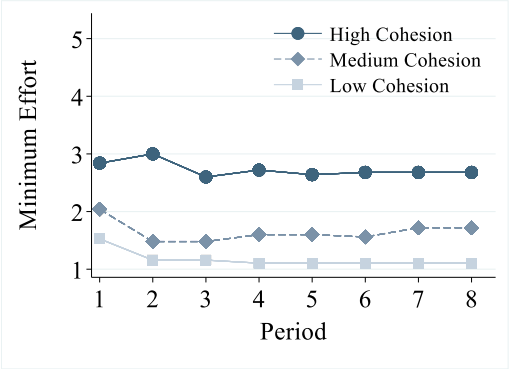
A pre-registered replication



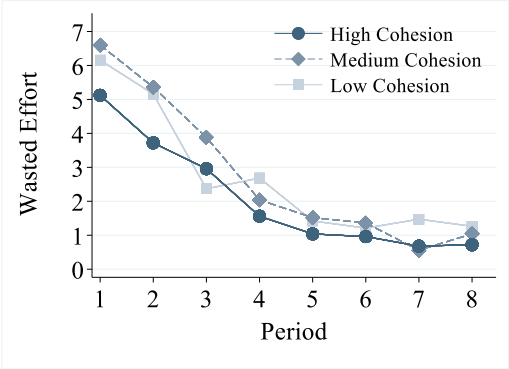
4a. The link between group cohesion and group-minimum effort in Period 1 and 8

[†††] Study 2 also closely replicates the evidence that the cohesion index varies coherently with tangible characteristic of the groups (See Fig. SM5.1 and Table SM2.1).

[‡‡‡] 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).



4b. Group minima across periods

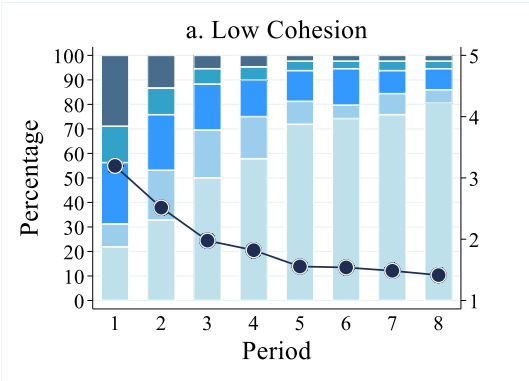


4c. Wasted effort

Figure 4. Study 2: Pre-registered replication independently conducted at University of Birmingham. Fig. 4a: Size of symbols proportional to the number of observations; in Periods 1 and 8, three N-matching observations are not displayed as they coincide with F-matching circles at coordinates (2, 1), (2.375, 2) and (2.875, 1); OLS Regression (69 groups), Period 1: $\beta = 0.321$ (se = 0.099, $p = 0.002$, $R^2 = 0.135$); Period 8: $\beta = 0.405$ (se = 0.108, $p < 0.001$, $R^2=0.175$); ordered probit estimation generates qualitatively similar results. Fig. 4b and 4c: "Low Cohesion" Partition (19 groups): group cohesion $\in$ [1, 2]; "Medium Cohesion" Partition (25 groups): group cohesion $\in$ (2, 4]; "High Cohesion" Partition (25 groups): group cohesion $\in$ (4, 7]. Fig. 4b: average group minimum effort over time. Fig. 4c: wasted effort per period is the sum of efforts in a group above the group minimum, averaged across groups.

C. Individual-Level Effort Choice

In this sub-section, we dig down to examine the association between *individual level* effort and group cohesion using pooled data from Studies 1 and 2 (see Fig. SM6.1, for corresponding analysis separately by study). Fig. 5 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.



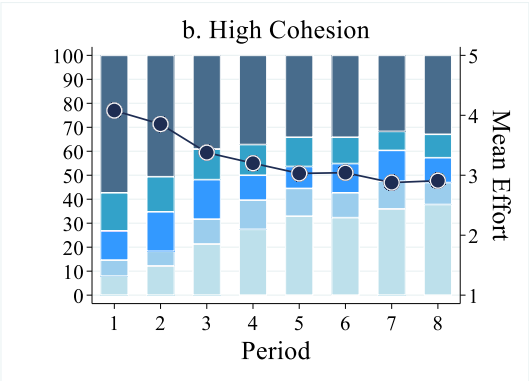


Figure 5. Study 1 and 2 combined: distribution of individual efforts over periods. Panel a: "Low Cohesion" Partition (32 groups): group cohesion $\in$ [1, 2]; Panel b: "High Cohesion" Partition (41 groups): group cohesion $\in$ (4, 7]. The bars represent the percentage of each effort level ranging from 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). Supplemental Material SM6 provides further analysis for all three partitions, separated by study.

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.[§§§]

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; *moderate* agents reduce effort but no lower than the previous period minimum; *harsh* agents reduce their effort below the previous minimum. We conjectured that subjects with high average oneness ratings of their fellow group members would be more likely to be nice, with the reverse true for individuals with low average oneness ratings of their fellow group members. An ordered probit analysis (using cases where a subject did not choose the minimum effort in period t) shows that their reaction in $t + 1$ (coded 1, 0 or -1 for nice, moderate, or harsh) varies positively with their average oneness ratings of the other three group members ($\beta = 0.102$, $p < 0.001$, pooled for Studies 1 and 2).

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 robustness and scale of that relationship through two sets of additional analyses.

A. Does Group Cohesion Outperform Homophily Measures as a Predictor of Effort?

This sub-section presents regression analysis assessing the power of group cohesion as a predictor of minimum effort with a particular focus on the impact of controls for homophily.

[§§§] In a nested random model (GLLAMM, Rabe-Hesketh et al., 2005) individual effort increases with the mean oneness rating of others in their group (β=0.139; p=0.001; Study 1 and 2 combined). Period dummies are negative ($p<0.01$); the oneness ratings' standard deviation is not significant. Ordered probit analysis (clustered on individuals) confirms these conclusions.

Via this analysis we address an issue raised in the introduction: since we interpret group cohesion as capturing the effects of real relationships that exist between group members, could we achieve comparable or better predictive power through use of information about observable individual characteristics? The main analysis we report makes use of the homophily index, first introduced in Section III, but here we provide more details of its construction.

Table 2 reports results for three models of group-level minimum effort which feature either group cohesion or the homophily index or both as independent variables. The homophily index combines data on 15 individual characteristics that we measured for this purpose (see SM1 for details).[****] For each of these 15 variables, we construct an homophily sub-index by first coding observations for each variable into a small number of mutually exclusive categories (e.g., two genders; three nationality groups). For each variable, we then assign a homophily sub-index to each group calculated as the proportion of group members associated with the highest-proportion category (e.g., suppose that, in a group, 3 members are female and 1 is male then, by definition, the gender homophily sub-index for that group is 3/4=0.75). The homophily index used in the regressions of Table 2 is then the average of the 15 sub-indices for each group. The models are estimated using standard ordered probit with clustering at group-level, since groups make multiple decisions.[††††] The regressions pool data for all 8 periods with separate panels for Study 1 (Panel A), Study 2 (Panel B) and the combined data set (Panel C).

[****] The 15 variables are: gender; age; field of study; nationality; no. of siblings; income; city size; no. of cohabitees; monthly budget; extent of self-finance; no. of club/group memberships; religiousness; political attitude; current happiness; future happiness.

[††††] We reach consistent conclusions if instead we account for interdependence of observations by estimating nested random models using GLLAMM (for details see Table SM7.1).

Table 2. Ordered Probit Regressions of Minimum Effort on Group Cohesion and Homophily

Panel A - Study 1						
Dep. var.: Min. Effort	(1)		(2)		(3)	
Group cohesion	0.448***	(0.105)			0.484***	(0.136)
Homophily index			3.871**	(1.809)	-1.038	(2.315)
Log-likelihood	-644.2		-681.6		-643.6	
# level 1 (resp. 2) units	520 (65)		520 (65)		520 (65)	
Panel B - Study 2: Pre-registered replication independently conducted at the BEEL Lab.						
Dep. var.: Min. Effort	(4)		(5)		(6)	
Group cohesion	0.388***	(0.097)			0.325**	(0.128)
Homophily index			6.400***	(2.468)	2.640	(2.883)
Log-likelihood	-569.0		-592.8		-565.3	
# level 1 (resp. 2) units	552 (69)		552 (69)		552 (69)	
Panel C - Study 1 and 2 combined						
Dep. var.: Min. Effort	(7)		(8)		(9)	
Group cohesion	0.414***	(0.074)			0.391***	(0.095)
Homophily index			3.240**	(1.486)	0.113	(1.803)
Study 2 (dummy var.)	-0.342*	(0.019)	-0.303	(0.195)	-0.351*	(0.191)
Log-likelihood	-1231.1		-1295.7		-1230.4	
# level 1 (resp. 2) units	1072 (134)		1072 (134)		1072 (134)	

Notes. Data from Periods 1 to 8. Variables are at group level. Variable definition and construction are in the supplemental material, section SM1. Period dummies (always included, relative to Period 1) are significantly negative (at p<0.05). Controls for individual effects: group-level clustering. Robust standard errors in parentheses. *** p<0.01, ** p <0.05 * p<0.1.

The estimated models that include group cohesion without homophily (models 1, 4 & 7 in Table 2) show that cohesion is a stable and strongly significant predictor across the two subject pools and the pooled data set. Similarly, when the homophily index enters without cohesion, consistent with our prior expectation, we find a strongly significant association for homophily in each case (models 2, 5 & 8). Critically, however, when both variables enter together, the homophily index is never significant while group cohesion remains strongly significant and with a coefficient very similar to that in the regression without the homophily index. As robustness checks, we conducted similar analysis along two further routes either entering all 15 homophily sub-indices as separate variables in regressions alongside group cohesion, or by using the 6 main principle components of the 15 homophily sub-indices as regressors alongside group cohesion (see SM8 for details).[‡‡‡‡] The consistent outcome of this analysis is that various

[‡‡‡‡] We are grateful to an anonymous referee for suggesting these robustness checks.

homophily-inspired measures do not match the performance of the group cohesion index in predicting group minimum effort.

B. Assessing the Magnitude of Cohesion-Related Cooperation

In this sub-section, we consider the *magnitude* of the effects of group cohesion on minimum effort, observed in our data. As one approach to this, 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. The results (presented in detail in SM9) demonstrate a very 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 cohesion (equal to 1): such a group is almost certain to be at minimum effort of 1 (the actual probability of minimum effort in this case is approximately 93 percent, based on pooled data from Studies 1 and 2). By contrast, a group with maximum possible group cohesion (equal to 7) is unlikely to end up at minimum effort of 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.

One might wonder how far these results depend on the 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 (see supplemental material, Table SM10.1).

As a second approach to assessing the scale of cohesion effects, 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). In these treatments, in line with the earlier research by Brandts and Cooper (2006) and others, we recruited unrelated individuals (not groups of friends) and they completed 8 rounds of the weak-link game. The bonus rates in four between-subjects' treatments (60 subjects each) were set at 6, 14, 22 and 30, respectively (see supplemental material, Table SM3.1, 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). Increasing the bonus monotonically increased the average minimum effort. At bonus level 6, it was 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, see Fig. SM12.1). Our results show 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 (see SM12 and Fig. SM12.1 for details). For example, 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 – or more precisely the value of the factors it proxies – is substantial, when gauged by the financial incentives needed to induce effort levels comparable to those of high cohesion groups.

VII. Towards an Explanation of the Power of Group Cohesion

Bringing real groups to the lab, as we have done, is a departure from classic lab experiments which might, initially, trouble 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. Yet working with real groups does create some methodological challenges and issues of interpretation, one of which we address next.

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, thus changing the payoff structure of the weak-link game making cooperation easier.[§§§§] In 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 game decisions. Our robustness tests extend the analysis of Table 2 by adding controls for self-reports of sharing. While this reached significance in some specifications, it had only a very modest impact on the coefficient for group cohesion which remained strongly significant in all cases (see SM11 for details). While this is reassuring, self-reports of sharing may not be entirely reliable and they may also be partly endogenous to game play.

With these limitations in mind, we ran a further set of treatments that we call the *Sharing Study* (Table 3, Appendix). For this study, we recruited fresh participants *individually* (hence, subjects typically did not know any other participant). Subjects played the weak-link game of

[§§§§] While the Nash equilibria are unchanged, the risks of cooperating are substantially reduced in groups committed to "full sharing" of payoffs, making cooperation easier to achieve.

Table 1 (where $b = 6$) following other standard procedures used across our studies but with the distinguishing feature that, before making their game decisions, 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. 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 any potential sharing arrangement.

The treatment where sharing is certain generated an average minimum effort of 2.73 which is comparable to the expected minimum effort associated with a group cohesion of close to 5 (see SM12). While introducing a little uncertainty about sharing (by setting the sharing probability at 80%) depressed average minimum effort a little (to a value just below 2.5), when the likelihood of sharing was only 50%, average minimum effort fell dramatically to a level only slightly above 1 (see Fig. SM12.1). While this evidence does not eliminate the possibility that expectations of sharing played some role, it counts against it being a convincing explanation of the broad patterns in our data: this is so because the ceiling of the sharing effect is well below the predicted effect of maximal cohesion (=7) and because uncertainty about sharing – quite likely in any actual sharing arrangements – rapidly diminishes its impact.

The results of the Sharing Study are interesting for the further reason that the treatment where sharing is certain can be interpreted as implementing an extreme form of social preferences in which each agent places the same weight on the earnings of all group members, including themselves. Viewed in this way, the results are consistent with some explanatory role for social preferences, albeit a limited one. A natural question is then, what is the relative importance of social preferences versus beliefs in mediating the impact of cohesion on effort?[*****]

We offer some tentative insight to this, exploiting data on beliefs and social preferences collected as part of Study 2. 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.[†††††] Then, at the end of the study, we measured participants' social preferences via a set of standard Social Value Orientation tasks: the "Social Value Orientation Slider Measure" due to Murphy,

[*****] In practice, it will be difficult to separate these roles clearly. For example, if groups with higher cohesion care more about each other's payoffs, in theory this reduces strategic risk, which in turn supports the expectation of higher effort levels within a group.

[†††††] In line with Schlag, Tremewan and van der Weele (2015), p. 484, we use 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 SM14.c, for details.

Ackermann and Handgraaf (2011).[‡‡‡‡‡] We use responses to these two sets of tasks as key inputs to a decomposition analysis based on the following simultaneous equation model:

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

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

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

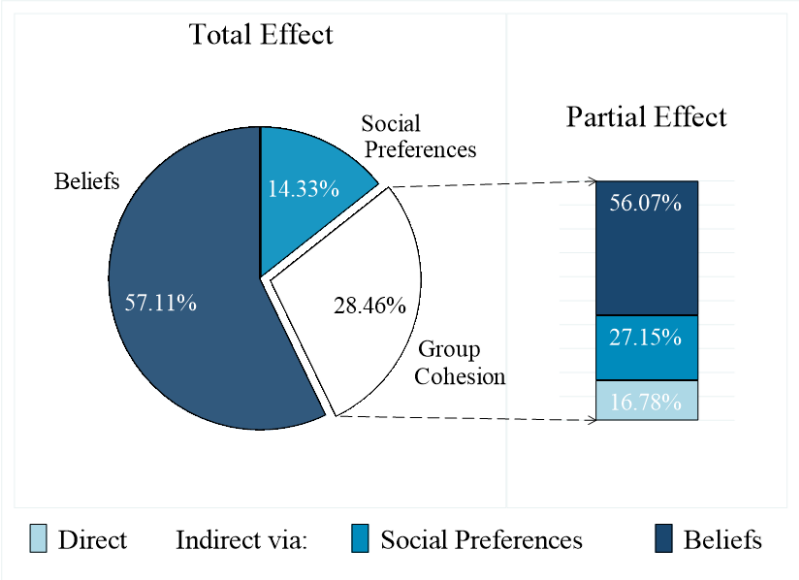
The first equation posits beliefs, (social) preferences and group cohesion as determinants of minimum effort. Group cohesion is treated as the unique (a priori) exogenous variable which can influence minimum effort directly (Eq. 1) and, indirectly, via beliefs (Eq. 2) or social preferences (Eq. 3).[§§§§§] 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 (Eq. 2). Although very simple from a psychological point of view, the model is presented in the spirit of a tool for assessing the relative importance of beliefs and social preferences as channels mediating the impact of group cohesion on effort, in our data.[******]

The estimated model produces significant coefficients (at 5% or 1% levels) for every β coefficient except one: specifically, we find no direct effect from social preferences to minimum effort (i.e., β_2 is not significantly different from zero). Hence, group cohesion impacts minimum effort through three active channels: it operates directly (via β_3) 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 SM13.1.

[‡‡‡‡‡] Each participant made 15 dictator style allocation decisions for an identified recipient from their group. The participant knew that one of the other two group members would make allocations to them (hence eliminating reciprocity considerations) but they did not know which one. See the supplemental material, section SM14.c, for further details.

[§§§§§] The analysis is conducted at group level. We use the average of the individual beliefs in each group and the average of the individual social value orientations in each group.

[******] The approach is similar in spirit to the mediation analysis reported in Kosse et al. (2020).



Decomposition of Total/Partial Effect on Minimum Effort

Figure 6. Study 2: modelling how group cohesion affects minimum effort. The panel reports the decomposition of the total/partial effect on minimum effort based on estimates derived from estimation of equations 1-3 above.

Figure 6 summarizes decomposition analysis conducted to assess the relative contributions of these three channels. While the pie chart provides a summary of the complete decomposition for the whole model, our primary interest is in the relative sizes of the partial effects (listed on the right-hand side of Fig. 6) which decompose the total effect of group cohesion into its three constituent paths. The path from group cohesion through beliefs accounts for about 56% of the total effect of group cohesion on minimum effort. While the impact via social preferences also accounts for a non-trivial proportion (about 27%) of the total effect, this path operates only indirectly via the beliefs channel suggesting that the role of social preferences is secondary to beliefs in both scale and mechanics (i.e., no direct effect of social preferences). Finally, the direct effect from group cohesion to minimum effort accounts for 16.7% of the total effect of group cohesion. We interpret the small size of this direct effect as "good news" in the sense that the impact of the factors proxied by group cohesion can be largely explained through its influence on the familiar rational choice concepts of beliefs and preferences.

For a variety of reasons, we suggest that the results of this decomposition be treated as tentative, absent further replication or other support. For example, we note a difference between the status of our measurements of beliefs and social preferences: specifically, while elicited beliefs measure something intrinsic to the weak-link games played by our participants, the measured social preferences capture something external to the game context. This might

have led to underestimation of the role of social preferences.[††††††] We could also measure social preferences in multiple different ways and an approach combining alternative ways of measuring them (à la Gillen, Snowberg and Yariv (2019)) could be an interesting avenue for checking the robustness of our conclusions from the mediation analysis. In addition, we cannot rule out the possibility that measured social preferences were to some extent influenced by experiences in play of the weak-link games although, conditional on there being such an effect, it seems most plausible to assume that success in the weak-link game would have encouraged more generous allocations in SVO tasks. In that case, our decomposition analysis should be interpreted as identifying an upper bound on the contribution of social preferences. Notwithstanding these potential reservations, however, the fact that the lion's share of the work is done by beliefs in our data stands in distinct contrast to results based on experiments using artificially-induced groups (see Chen and Chen (2011)). At minimum, we therefore suggest that our results should unsettle any presumption that social preferences are the primary channel through which within-group relationships affect success in team production.

VIII. Conclusions

It is hard to deny that social relationships may 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, adopting the weak-link setting as a workhorse for proof of concept. 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, et al. (2015)). We used our measure of group cohesion, which is a 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 1160 participants and including a variety of robustness tests, benchmarking exercises, and an independent pre-registered replication, we

[††††††] We are grateful to an anonymous referee for highlighting this possibility.

examined the predictive power of group cohesion in the context of experimental weak-link 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 low cohesion groups rarely, if ever, do well. We used an econometric approach to explore possible mechanisms underpinning the association between group cohesion and group minimum effort and found that, in our model, group cohesion shapes both beliefs and social preferences but with beliefs emerging as the primary channel. We have also presented evidence that the changes in effort associated with variation in cohesion can be considered "large" in the context we have studied.

While we cannot directly extrapolate to predict the scale of comparable effects in other lab or in field contexts beyond those we have studied, our results do provide motivation for exploring such issues using our group cohesion index. On the assumption that our results do translate to the field, they have particular potential significance in the context of organizational performance (e.g., Akerlof and Kranton (2005); Ashraf and Bandiera (2018)). If group cohesion is associated with desirable team or group outcomes across a variety of organizational settings, then our tool may 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, beyond the new evidence we have presented, we believe we have provided proof of concept for a new simple and portable tool designed to facilitate the quantitative study of social relationships as factors of team production.

Data availability

Data and analysis code are available at https://osf.io/g9u3e.

REFERENCES

Akerlof, George A., "Social Distance and Social Decisions," *Econometrica*, 65 (1997), 1005-27.

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

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

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

Babcock, Philip, Kelly Bedard, Stefanie Fischer, and John Hartman, "Coordination and Contagion: Individual Connections and Peer Mechanisms in a Randomized Field Experiment," *Journal of Public Economics*, (2019), 104069.

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

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

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

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

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

Brandts, Jordi, and David J. Cooper, "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 (2007), 1223-68.

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

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

Charness, Gary, and Yan Chen, "Social Identity, Group Behavior, and Teams," *Annual Review of Economics*, 12 (2020), 691-713.

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

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

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

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

Chierchia, Gabriele, Fabio Tufano, and Giorgio Coricelli, "The differential impact of friendship on cooperative and competitive coordination, " *Theory and Decision* 89 (2020), 423-452.

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

Cooper, David J., and Roberto A. Weber, "Recent Advances in Experimental Coordination Games," in Monica Capra, Rachel Croson, Mary Rigdon and Tanya Rosenblatt ed., *The Handbook of Experimental Game Theory*, (Cheltenham: Edward Elgar Publishing, 2020).

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

Currarini, Sergio, and Friederike Mengel, "Identity, Homophily and in-Group Bias," *European Economic Review*, 90 (2016), 40-55.

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

Dunbar, R. I. M., "The Anatomy of Friendship," *Trends in Cognitive Sciences*, 22 (2018), 32-51.

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

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

Gächter, Simon, Chris Starmer, Christian Thöni, Fabio Tufano, and Till O. Weber, "Social closeness can help, harm and be irrelevant in solving pure coordination problems, " *Economics Letters*, 216 (2022), 110552.

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

Gillen, Ben, Erik Snowberg, and Leeat Yariv, "Experimenting with Measurement Error: Techniques with Applications to the Caltech Cohort Study," *Journal of Political Economy*, 127 (2019), 1826-63.

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

Glaeser, Edward L., David Laibson, and Bruce Sacerdote, "An Economic Approach to Social Capital," *Economic Journal*, 112 (2002), F437-F58.

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

Goyal, Sanjeev, "Strong and Weak Links," *Journal of the European Economic Association*, 3 (2005), 608-16.

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

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

Jackson, Matthew O., "A Typology of Social Capital and Associated Network Measures," *Social Choice and Welfare*, 54 (2020), 311-36.

Kelley, Harold H. , Ellen Berscheid, Andrew Christensen, John H. Harvey, Ted L. Huston, et al., *Close Relationships*, (New York: Freeman, 1983).

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

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk, "The Formation of Prosociality: Causal Evidence on the Role of Social Environment," *Journal of Political Economy*, 128 (2020), 434-67.

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

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

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

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

Putnam, Robert, *Bowling Alone. The Collapse and Revival of American Community*, (New York: Simon & Schuster, 2000).

Quidt, Jonathan de, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and Bounding Experimenter Demand." *American Economic Review* 108 (11): 3266–3302.

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

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

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

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

van Huyck, John B., Raymond C. Battalio, and Richard O. Beil, "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review.*, 80 (1990), 234-48.

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

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

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

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

APPENDIX

Table 3. List of Experimental Studies

		Research Objectives	Incentive Structure	Recruitment	Random Allocation	N
I.	Study 1	Predictive Power of Group Cohesion	b=6	Friends	F- or N-matching	260
II.	Two-week Study	Construct reliability (Test-retest reliability; Task sequencing)	b=6	Friends	F- or N-matching	172
III.	Study 2	Replicating Study 1; Mediational channels: Beliefs, Social Prefs.	b=6	Friends	F- or N-matching	276
IV.	50-period Study	Long horizon	b=6	Strangers	Groups	32
V.	Share Study	To compare the cooperation enhancing effects of group cohesion with sharing rules	b=6, Pr{*S*}=0.5	Strangers	Groups	60
			b=6, Pr{*S*}=0.8	Strangers	Groups	60
			b=6, Pr{*S*}=1	Strangers	Groups	60
VI.	Bonus Study	To compare the cooperation enhancing effects of group cohesion with financial bonuses	b=6	Strangers	Groups	60
			b=14	Strangers	Groups	60
			b=22	Strangers	Groups	60
			b=30	Strangers	Groups	60

Note. Study 2 was a pre-registered (see footnote 7) replication independently conducted at the BEEL lab (University of Birmingham, UK) by in-house experimenters. All the other studies were conducted at the CeDEx Lab (University of Nottingham, UK). Total overall sample: 1160 participants. *b* is the bonus rate controlling the marginal return to changes in minimum effort. Pr{*S*} stands for probability of sharing.

Online Supplemental Material (SM) for

Measuring "Group Cohesion" to Reveal the Power of Social Relationships in Team Production

Simon Gächter[1,2,3], Chris Starmer[1,*] and Fabio Tufano[1,*].

[1] CeDEx, School of Economics, University of Nottingham, NG7 2RD
[2] IZA, Bonn, Germany
3 CESifo, Munich, Germany
* Corresponding authors (chris.starmer@nottingham.ac.uk, fabio.tufano@nottingham.ac.uk).

25 July 2022

Table of Contents

SM1: Definitions of Variables used in Homophily Index Construction

In this section, we define 15 individual-level variables that feature in the analysis related to homophily. These draw on data from a subset of the characteristics elicited by the control tasks at the end of each experimental session (see column 1 of Table SM1.1 for the list of characteristics). Table SM1.1 also explains the construction of the corresponding group-level *homophily sub-index* (see column 3) for each of the 15 individual-level variables listed in column 2. What we refer to as the *homophily index* in the main text is computed as the overall average of the fifteen group-level homophily sub-indices, constructed as per colum 3 of Table SM1.1.

TABLE SM1.1

DEFINITIONS OF INDIVIDUAL-LEVEL VARIABLES AND CONSTRUCTION OF HOMOPHILY INDICES

#	**Elicited Characteristic** (1)	**Individual-level Variable and its characteristics** (2)	**Group-level Homophily Sub-Index (Construction)** (3)
1	*Nationality*	Categorical variable assuming values 1 (British), 2 (EU), and 3 (other). [3.]	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 above categories.
2	*Age*	An interval variable measured in years. [2.]	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 of studies*	Categorical variable taking values 0 (arts and education), 1 (economics and business), 2 (social sciences), 3 (medicine), and 4 (engineering and natural sciences). [5.]	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	*Gender*	Dummy variable equal to 1 for "female" and 0 for "male". [1.]	From the two mutually exclusive categories (i.) Female; (ii.) Male, calculate the highest proportion in each group of members from one of the two above categories.
5	*Religiousness*	Indicates individual religiosity and ranges from 1 (not at all religious) to 7 (very religious). [12.]	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 5 on a Likert scale from 1 "Not at all religious" to 7 "Very religious"; (iii.) more than 5 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.
6	*No. of club memberships*	Counts the number of memberships in six possible categories of voluntary associations. It takes values from 1 (no membership at	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

		all) to 7 (membership in all six categories). [11.]	group of members from one of the three above categories.
7	*Monthly budget*	Reports (in pounds) the budget available per month (expenses for accommodation already detracted). [9.]	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.
8	*No. of siblings*	Interval variable documenting the number of siblings. [4.]	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.
9	*Cohabitee*	Interval variable indicating the number of people in the household (respondent included). [8.]	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.
10	*City size*	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.]	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.
11	*Self-financed*	Indicates the percentage of the monthly expenses that is self-financed. [10.]	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.
12	*Political attitude*	Describes political alignment by ranging from 1 "left-wing" to 7 "right-wing". [13.]	First, generate 3 mutually exclusive categories: (i.) up to 2 on a Likert scale from 1 "Left" to 7 "Right"; (ii.) between 2 and 5 on a Likert scale from 1 "Left" to 7 "Right"; (iii.) more than 5 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.
13	*Income rank*	Categorical variable for the relative family income of the respondents, ranging from 1 (far below average) to 5 (far above average). [6.]	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.
14	*Future happiness*	Categorical variable documenting the life satisfaction expected in five years' time. It varies from 1 (not at all satisfied) to 10 (absolutely satisfied). [15.]	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	*Current happiness*	Categorical variable indicating current life satisfaction. It takes values from 1 (not at all satisfied) to 10 (absolutely satisfied). [14.]	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.

Note.— In column 2, the number in square brackets reports the reference number of the relevant question in section SM14.d. Not all the variables measured in the post-experimental survey feature in analysis reported in the paper or supplementary materials. Those that do are described in SM1 and SM11. The full survey is described in SM14. Variables measured in the survey but not used in reported analysis fall mainly into two categories. One category includes variables collected in the expectation that they might have significant explanatory power in models of group minimum effort but which we dropped from the reported analysis due to lack of significance: this includes measures of risk attitude (elicited as Questions 16-21) and measures related to generalized trust (Questions 22-25) in SM14.d below. A second class of variables includes those collected for low-ball tests of construct validity (less sophisticated, for instance, than the analysis of homophily reported in the paper): this includes Questions 26 and 27 (see SM14.d). We considered these to be capturing aspects of group cohesion and as expected, they were positively correlated with it (but considerably less successful in predicting group minimum effort).

SM2: Study 1 and 2: How Do Individual Characteristics Vary Within and Between Groups?

A key design objective was to import "real" groups into the lab. From the sociological literature on friendship we know that people tend to become friends with people who share similar socio-demographic backgrounds, attitudes and preferences (e.g., McPherson, Smith-Lovin and Cook (2001)): that is, "like-befriends-like". Hence, if we have succeeded in this objective, for F-matching groups, we should find greater variation of socio-demographic and attitudinal characteristics *between* groups than *within* groups. We tested this prediction separately for each of 15 individual characteristics defined in Table SM1.1 above using both ANOVA and Kruskal-Wallis tests. Table SM2.1 reports test results for both Study 1 (see Section IV in the main paper) and Study 2 (a pre-registered replication of Study 1, see Section V.B in the main paper).

In line with our prediction, based on the non-parametric Kruskal-Wallis test, out of the 15 characteristics the null of equal variance between and within groups is rejected at $p < 0.05$ for 11 characteristics in Study 1 and 13 characteristics in Study 2, respectively. By contrast, and as expected, corresponding analysis for N-matching groups reveals almost no significant differences for any characteristic (with the only exception of *current happiness* which is marginally significant at the 10% level in Study 1), comparing within and between group homogeneity in either Study 1 or Study 2. Similar inferences can be drawn based on the parametric ANOVA test results in Table SM2.1.

TABLE SM2.1

STUDY 1 AND 2: TESTING FOR EQUALITY OF VARIANCE WITHIN AND BETWEEN GROUPS (F-MATCHING)

		Study 1			Study 2		
	Individual-level	ANOVA		K.-W. test	ANOVA		K.-W. test
#	Variable	F	R^2	χ^2 (with ties)	F	R^2	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	Religiousness^i	2.630***	0.462	82.772***	4.28***	0.583	109.375***
6	No. of club memberships^i	2.350***	0.434	80.388***	1.920***	0.385	69.946**
7	Monthly budget^i	1.056**	0.337	74.820***	2.390***	0.439	92.216***
8	No. of siblings^i	2.020***	0.397	71.930***	2.160***	0.413	88.313***
9	Cohabitee^i	2.190***	0.417	69.942**	1.810***	0.371	81.649***
10	City size^i	1.750***	0.364	68.677**	1.520**	0.332	65.152*
11	Self-financed^i	1.550**	0.336	63.338**	1.290	0.297	60.666
12	Political attitude^i	1.360**	0.308	60.991*	1.970***	0.392	81.244***
13	Income rank^i	1.280	0.295	60.255*	1.830***	0.374	69.413**
14	Future happiness^i	1.020	0.249	49.863	1.610**	0.344	61.682*
15	Current happiness^i	1.060	0.257	47.546	1.380*	0.311	63.135

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 above in Table SM1.1. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

SM3: Bonus Levels, Payoff Matrices and the Harshness of the Game Environment

Table SM3.1 illustrates the payoff matrices used in the four treatments of the Bonus Study (see Table 3 in the paper's appendix, for key details of this study). In each treatment, 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 the Bonus Study, we set b equal to either 6, 14, 22 or 30, respectively, in one of the four treatments. (In all other studies we report in the paper, b=6).

TABLE SM3.1

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

Bonus $b = 6$						*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 $b = 14$						*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 $b = 22$						*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 $b = 30$						*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, defined below.

The bonus *b* can be interpreted as controlling the "harshness" of the weak-link game. Loosely speaking, the game may be considered harsher, the more difficult it is for groups to achieve coordination on the better equilibria. Lower bonuses imply harsher environments and, with the bonus set as 6 (as it was in our Studies 1 and 2), we expected that low cohesion groups would typically find

it difficult to escape the Pareto-worst equilibria. This expectation was based on evidence from a series of papers in which Jordi Brandts, David Cooper and co-authors used a similarly calibrated game with $b=6$ – the inspiration for our base game – to "trap" experimental groups in the lowest Pareto-ranked equilibrium before implementing their relevant experimental interventions (e.g., Brandts and Cooper (2006); Brandts and Cooper (2007); Brandts, Cooper and Fatas (2007); Brandts, Cooper and Weber (2015)).

The concept of harshness can be formalized through an index $h=(mr/n)/mc$, where mr is the marginal return to minimum effort; n is the number of players and mc is the individual marginal cost (see Anderson, Goeree and Holt (2001); Chen and Chen (2011)). The lower h is, the harsher the game. Our studies with $b=6$ imply $h=0.30$ (i.e., $mr/n=(10\cdot b)/4=60/4=15$ while $mc=50$, hence $h=0.30$). As a comparator benchmark, in Chen and Chen (2011) $h=0.67$ (i.e., a less harsh environment).

Table SM3.1 reports the harshness index h for each of the four bonus treatments: look to the top right above each payoff matrix. In line with the above discussion, h increases with the bonus indicating that higher bonus treatments are less harsh. Consistent with this interpretation of falling harshness, we would expect groups to be more successful in achieving more highly Pareto-ranked equilibria in treatments with higher bonuses and this is exactly the tendency we find (see Section SM12 below).

SM4: Two-week Study: Additional Analysis

Fig. SM4.1 plots the association between average group cohesion and average minimum effort across Study 1 and the Two-week Study, after partitioning the data in low cohesion (i.e., *group cohesion* $\in$ [1, 2]), medium cohesion (i.e., *group cohesion* $\in$ (2, 4]) and high cohesion (i.e., *group cohesion* $\in$ (4, 7]) partitions.

A key feature of the Two-week Study is that we elicited the oneness ratings a week prior to the play of the weak-link game. Inspection of Fig. SM4.1 suggests that the prior elicitation of oneness may have affected the behavior of medium cohesion groups. We conjecture the following explanation. The fact that low or high cohesion groups seem unaffected 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 medium 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.

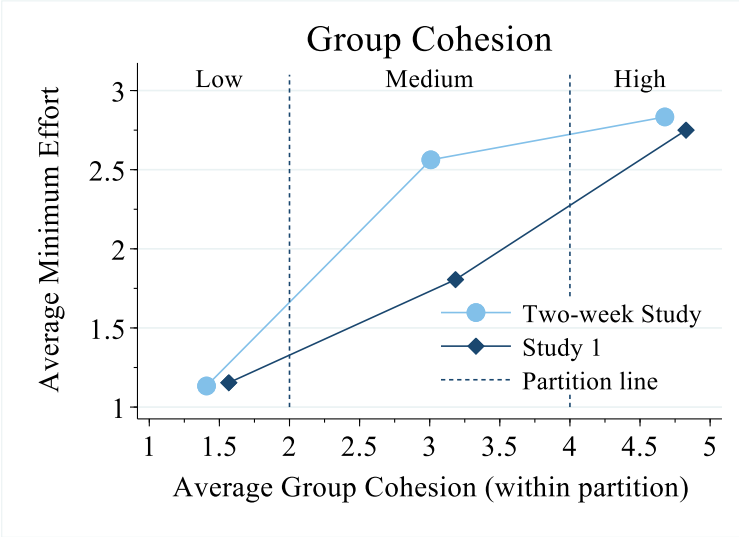


FIG. SM4.1. — Minimum Effort and Group Cohesion in the Two-week Experiment and Study 1 (Group-level data from Period 8). For comparability, the average group cohesion for the Two-week Study is based on week 2 group cohesion. "Low Cohesion" partition: *group cohesion* ∈ [1, 2]; "Medium Cohesion" partition: *group cohesion* ∈ (2, 4]; "High Cohesion" partition: *group cohesion* ∈ (4, 7].

In the light of the visual difference in behavior for medium cohesion groups between Study 1 and the Two-week Study, we proceed by checking the robustness of the relationship between group cohesion and minimum effort in the Two-week Study also to the exclusion of medium cohesion groups.[1] Table SM4.1 below reports results from OLS regressions of minimum effort on group cohesion (using data only from Period 8). The top row of the "Full Sample" column corresponds with the analysis presented graphically in Figure 3a (right-hand panel, Period 8) in our main paper. The bottom row of the same column shows the corresponding statistics calculated for the Two-week Study. Although slightly lower, the estimated coefficient is qualitatively similar in magnitude and the coefficients are both highly significant ($p < 0.01$). The robustness check runs the same pair of regressions but excluding observations for the medium-cohesion groups. The results are reported in the right-hand column of Table SM4.1 and it is apparent that the results are very similar when compared to the corresponding full sample results.

We expected that the relationship between group cohesion and minimum effort would be robust: this expectation was based partly on the fact that the means for group minimum effort are statistically indistinguishable across the two studies when comparing either the two means for the high partition or the two means for the low partition (using either a t-test or a Mann-Whitney test). Moreover, we only find weak statistical support for the conclusion that average minimum efforts differ between the two studies in the medium cohesion partition: the difference is not significant at any conventional level based on the Mann-Whitney test; it is just significant at the 10% level ($p=0.09$) using a 2-tailed t-test.

[1] We are grateful to an anonymous referee for suggesting these robustness checks.

TABLE SM4.1

TWO-WEEK STUDY: OLS REGRESSIONS OF MINIMUM EFFORT ON GROUP COHESION

	Group Cohesion Coefficients	
Period 8 data	Full sample	Excluding mid-range cohesion partitition
Study 1	0.55 (0.000)	0.54 (0.001)
Two-week Study	0.46 (0.005)	0.45 (0.008)

NOTE.— p-values in parentheses.

SM5: The Distribution of Group Cohesion in Study 2

Here we present the distribution of group cohesion for Study 2 via Fig. SM5.1 (as per the presentation of corresponding data for Study-1 in Fig. 2). There is substantial cross-group variation in group cohesion with scores ranging from 1 to 6.3 (for Study 1, scores ranged from 1 to 5.5). Within the F-matching, the range is between 1.6 and 6.3 (the Study-1 range was from 2.0 to 5.5), whereas in N-matching it is between 1.0 and 2.9 (compared with the range 1.0 to 3.3 for Study 1). Like Study 1, for F-matching groups in Study 2, the variation of oneness ratings is larger between groups than within groups (Kruskal-Wallis test: $\chi^2(46) = 101.7$, $p < 0.001$), whereas this is not the case in N-matching groups ($\chi^2(17) = 14.6$, $p = 0.751$). This confirms that Study 2 closely replicated the Study-1 findings on the distributions of group cohesion scores between N- and F-matching and, because of these characteristics, like Study 1, it provides good scope for examining how group cohesion and game behavior are linked.

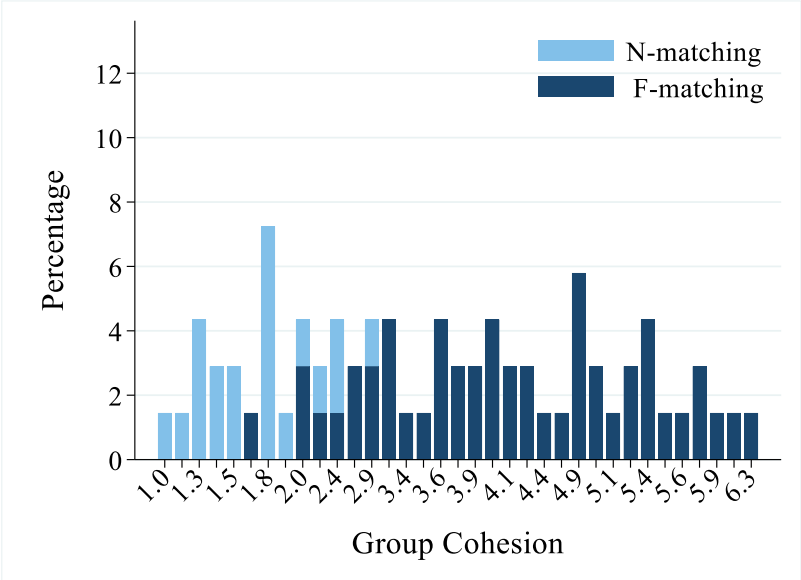


FIG. SM5.1.—Study 2: The distribution of group cohesion under F- and N-matching.

SM6: The Distribution of Individual Effort over Periods

In this section, we complement Fig. 5 of the paper by showing the distributions of individual effort in groups with, respectively, low, medium and high group cohesion, separately for Studies 1 and 2. For each panel in Fig. SM6.1, color coding shows the distribution of efforts while the average of individual effort is indicated with a circle. The panels *b* and *e* provide a visualization for the mid-range cohesion partition (not displayed in the main paper). From eyeballing panels *a*-*f*, it is evident that the time profile of average individual effort is very similar between studies (when comparing left- and right-hand panels for the same cohesion partition).

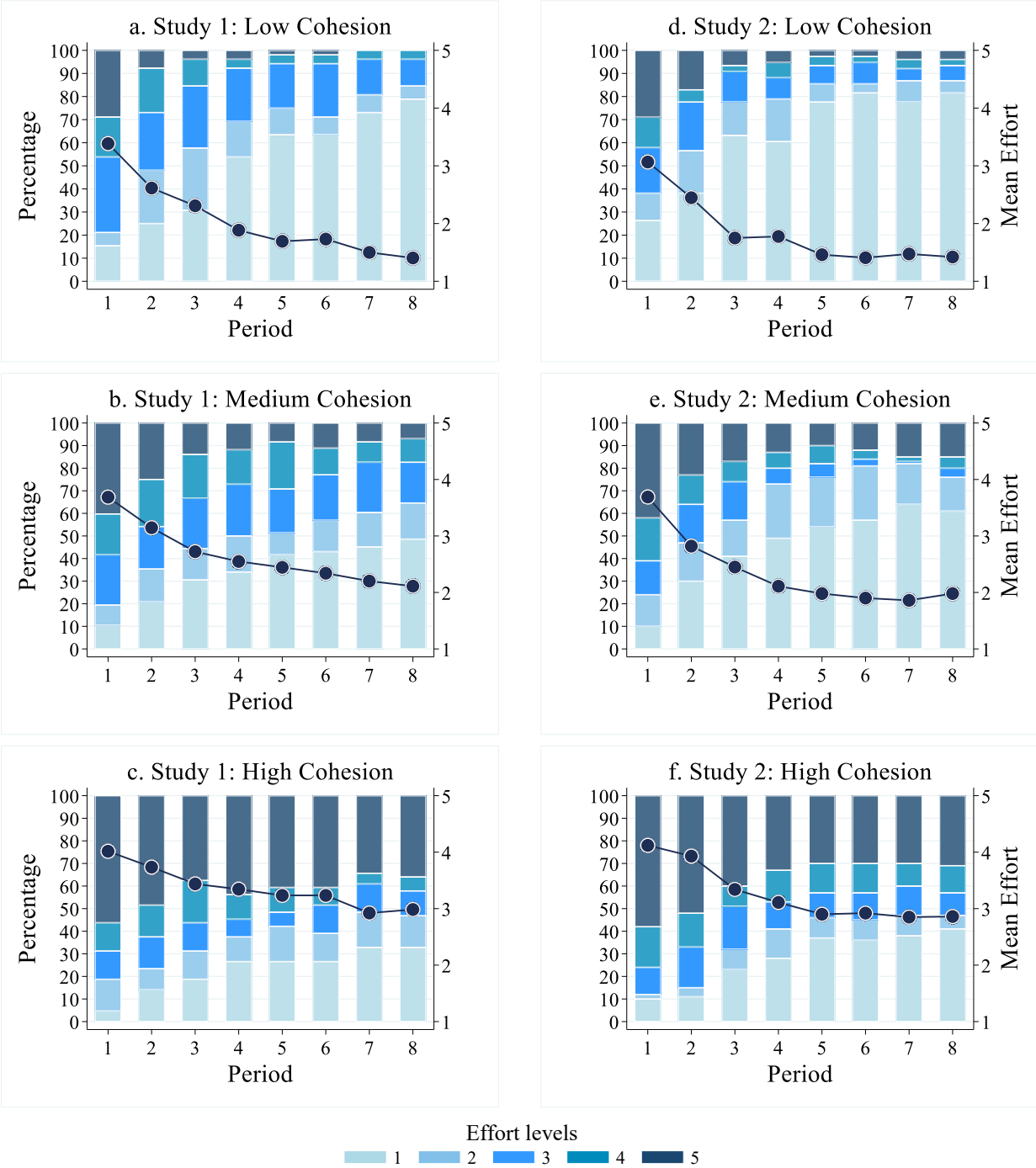


FIG. SM6.1.—Distribution of individual efforts over periods in Study 1 and 2. "Low Cohesion" Partition: group cohesion ∈ [1, 2]; "Medium Cohesion" Partition: group cohesion ∈ (2, 4]; "High Cohesion" Partition: group cohesion ∈ (4, 7]. Panel *a*: 13 groups; panel *b*: 36 groups; panel *c*: 16 groups; panel *d*: 19 groups; panel *e*: 25 groups; panel *f*: 25 groups. 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 individual level mean efforts (measured on the secondary y-axes).

SM7: Complementing Our Econometric Estimations: GLLAMM Regressions

TABLE SM7.1

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)
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 definitions: the construction of the Group cohesion index is explained in the paper (Section III.A); see SM1 for the Homophily Index and SM11 for Share. 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. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

To complement our econometric analysis reported in Table 2 in our paper and below in Table SM11.1, we take a different approach to accounting 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 Table SM7.1. The use of this estimation procedure along with the standard ordered probit regressions in Table 2 and Table SM11.1 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 prone to type II errors, while the more powerful nested procedure is prone to type I errors. Our main results are consistent across the two estimation procedures: group cohesion (conditional on being entered in the regression model) is always significant even in the presence of controls for homophily and share.

SM8: Econometric Specifications and Homophily

In the main paper, we show that group cohesion is a strongly significant predictor of group minimum effort (see Table 2) and that this remains so in the presence of an homophily index as a control. In this section, pooling data from Study 1 and 2, we explore the robustness of this result to two different ways of capturing the effects of homophily in the econometric specifications. In one case we use 15 separate homophily sub-indices in place of the aggregate index (see Table SM1.1 for definitions); in the other we use principal components of homophily sub-indices as controls. To preview the findings of these robustness checks: our main conclusions go through with either approach.

Analysis with 15 separate homophily sub-indices: Regression (1) in Table SM8.1 below reports an ordered probit regression of minimum effort on group cohesion along with the 15 homophily sub-indices and the Study 2 (dummy) variable. Only group cohesion is statistically significant (both with $p<0.01$), while none of the 15 homophily sub-indices is.

Principal component analysis: this analysis allows us to reduce the dimensionality of our homophily data by identifying the components which explain a significant fraction of the variance across the 15 homophily sub-indices. In accordance with the Kaiser criterion (which drops any component with eigenvalues smaller than 1) and the Scree test (involving a plot of the eigenvalues in decreasing order of their magnitude against the component numbers to determine where the eigenvalues level off), we retain only the first six components (cumulatively accounting for 59.77% of the variance), with the first component having an eigenvalue of 2.570 (accounting for 17.13% of

the variance) and the sixth component having an eigenvalue of 1.026 (accounting for 6.84% of the variance; the next biggest eigenvalue is 0.972 accounting for 6.48% of the variance).

TABLE SM8.1

STUDY 1 AND 2 COMBINED: ROBUSTNESS TESTS AND HOMOPHILY

Ordered probit regression of min. effort on the 15 homophily sub-indices			Ordered probit regression of min. effort on the homophily sub-indices principal components		
Dep. variable: Min. effort	(1)		Dep. variable: Min. effort	(2)	
Group cohesion	0.369***	(0.108)	Group cohesion	0.385***	(0.098)
Nationality	0.241	(0.670)	Principal Component 1	0.059	(0.086)
Gender	0.463	(0.573)	Principal Component 2	-0.042	(0.083)
Age	-0.238	(0.524)	Principal Component 3	-0.084	(0.092)
No. of siblings	-0.047	(0.472)	Principal Component 4	0.052	(0.096)
Field study	-0.159	(0.503)	Principal Component 5	0.150*	(0.084)
Income rank	-0.259	(0.540)	Principal Component 6	-0.047	(0.111)
City size	-0.127	(0.626)			
Cohabitee	0.125	(0.521)			
Monthly budget	-0.287	(0.593)			
Self-financed	0.950	(0.589)			
No. of memberships	0.163	(0.564)			
Religiousness	-0.009	(0.587)			
Political attitude	-0.537	(0.531)			
Current happiness	0.0298	(0.574)			
Future happiness	0.502	(0.493)			
Study 2 (dummy variable)	-0.272	(0.218)	Study 2 (dummy variable)	-0.342	(0.191)
Log-likelihood	-1208.575		Log-likelihood	-1215.392	
# level 1 (resp. 2) units	1072 (134)		# level 1 (resp. 2) units	1072 (134)	

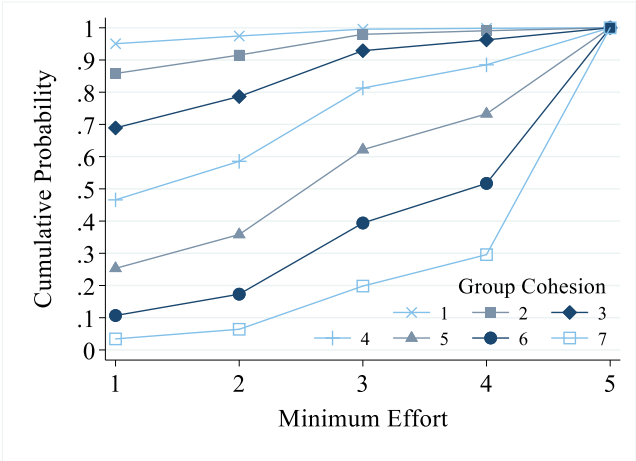
NOTES.—Data from Periods 1 to 8. Explanatory variables are at group level. Variable definitions: the construction of the group cohesion index is explained in the paper (Section III.A); see SM1 for definitions of the 15 homophily relevant variables. Period dummies (always included, relative to Period 1) are significantly negative (at $p < 0.01$). Controls for individual effects: clustering. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

Regression (2) in Table SM8.1 reports an ordered probit regression of minimum effort on group cohesion along with the 6 principal components and the Study 2 (dummy) variable. Group cohesion is strongly statistically significant. Of the other variables, just one principal component reaches significance (but only at the 10 percent level) and no other variable reaches significance at any conventional threshold of statistical significance.

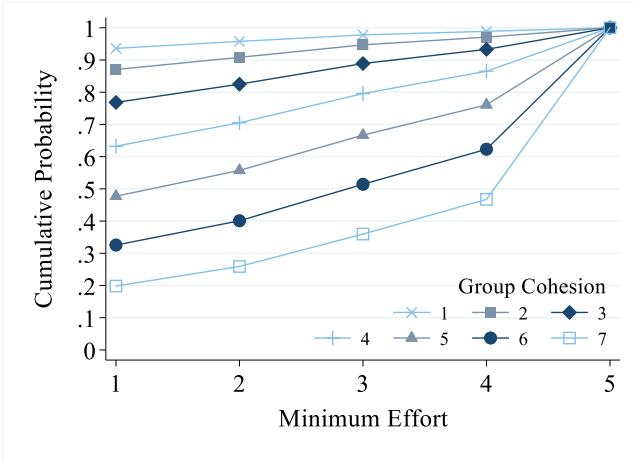
To sum up, both forms of analysis presented above confirm the robustness of our results: the estimated effects of group cohesion are not sensitive to the way in which homophily measures enter the analysis.

SM9: Supplementary Analysis on the Predictive Power of Group Cohesion

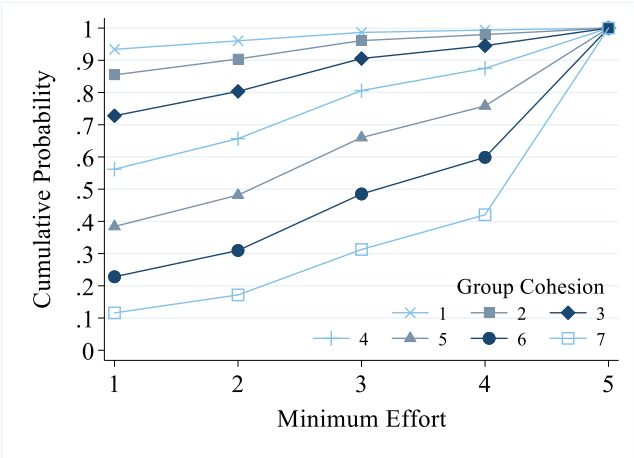
In this section we provide further analysis of the association between group cohesion and minimum effort by regressing minimum effort in the last period (period 8) on group cohesion to generate the predicted probabilities for each possible level of minimum effort, conditional on different levels of group cohesion. Fig. SM9.1 presents the results as cumulative distribution functions of minimum effort, with panel *a* using data from Study 1, panel *b* from Study 2 and panel *c* from Studies 1 and 2 combined. While displaying some minor differences, the predicted CDFs across all panels show a sizeable predictive power of group cohesion on minimum effort when going from the minimum to the maximum point on the group cohesion scale. For brevity, we illustrate with panel *c* (pooled data) but broadly similar patterns are evident in the corresponding analysis for the subsamples associated with Studies 1 and 2 (panels *a* and *b* respectively). To interpret Fig. SM9.1*c*, consider the same case discussed in the main text: that is, a group characterized by minimum group cohesion (equal to 1). From the top line plotted in panel *c* in Fig. SM9.1, we see that such a group is almost certain to be at minimum effort (the actual probability of minimum effort in this case is approximately 93 percent). By contrast, a group with maximum possible group cohesion (equal to 7 and represented visually by the bottom line plotted in the panel) is unlikely to end up at minimum effort of 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.



a. Study 1



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



c. Study 1 and 2 combined

FIG. SM9.1.—Predicted CDFs for minimum effort for each level of group cohesion (group-level data from period 8)

SM10: An Alternative Cohesion Metric: The Group Average of Individual Oneness Ratings

The results in Table 2 in our paper and the CDFs in Fig. SM9.1 point to a strong predictive power of group cohesion for 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, as shown in Table SM10.1, using the group average of individual oneness ratings as an alternative cohesion metric delivers very similar results to those detailed in our paper (Table 2). Relative to Table 2, in some of the regressions reported here we also include an extra control variable, "Share". The rationale for this extra control and its definition are provided in SM11.

TABLE SM10.1

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 (134)

NOTE.—Data from Periods 1 to 8. Explanatory variables are at group level. See SM1 for definition of the homophily index and SM11 for Share. 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. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

SM11: Econometric Models with Share as Control Variable

SM11.a Definition of the Share Variable

We use the label "Share" to refer to a variable based on the principal component of individual level responses to two tasks – *Share intention* and *Share influence* – which were elicited as part of the set of control tasks presented at the end of experimental sessions.

- *Share intention* is the self-reported intention to share the earnings with other group members after the experiment. It ranges from 1 (certainly not) to 7 (certainly).
- *Share influence* is the self-report of how the possibility of sharing experimental earnings afterwards affected decisions. It ranges from 1 (certainly not) to 7 (certainly).

The precise questions used to elicit these variables can be found in the experimental materials (see SM14) where they feature as Questions 29 and 30, respectively in the sequence of control tasks listed in SM14.d. These two questions are based on tasks used by Reuben and van Winden (2008) and we are grateful to Frans van Winden for suggesting we use them as part of our design.

SM11.b Analysis Exploiting Share as a Control

A possible interpretation of the relationship we observe 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.[2] 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. One dimension of this is the analysis we report in this section using the variable *Share* (defined as in SM11.a) as an additional control in our analysis of the association between group cohesion and group minimum effort (but see also Section VII of the paper and Section SM12 below for further analysis and commentary bearing on this general issue).

[2] Although 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 (defined in SM3 above) 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.

TABLE SM11.1

ORDERED PROBIT REGRESSIONS WITH SHARE AS CONTROL VARIABLE

Panel A – Study 1				
Dep. variable: Min. Effort	(1)	(2)	(3)	(4)
Group cohesion		0.389***		0.452***
		(0.123)		(0.143)
Homophily index			2.252	-2.032
			(0.233)	(2.292)
Share (principal component)	0.277**	0.164	0.233*	0.184
	(0.116)	(0.129)	(0.126)	(0.130)
Log-likelihood	-670.5	-639.7	-666.7	-634.6
# 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: Min. Effort	(5)	(6)	(7)	(8)
Group cohesion		0.300***		0.251*
		(0.108)		(0.133)
Homophily index			4.645*	2.254
			(2.559)	(2.845)
Share (principal component)	0.370***	0.221*	0.298***	0.211*
	(0.095)	(0.114)	(0.101)	(0.111)
Log-likelihood	-583.1	-557.5	-569.2	-554.9
# level 1 (resp. 2) units	552 (69)	552 (69)	552 (69)	552 (69)

Panel C – Study 1 and 2 combined				
Dep. variable: Min. Effort	(9)	(10)	(11)	(12)
Group cohesion		0.338***		0.335***
		(0.083)		(0.100)
Homophily index			3.240**	0.113
			(1.575)	(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)	-0.360*	-0.414**	-0.398**	-0.415**
	(0.193)	(0.195)	(0.196)	(0.195)
Log-likelihood	-1270.7	-1211.6	-1255.6	-1211.6
# level 1 (resp. 2) units	1072 (134)	1072 (134)	1072 (134)	1072 (134)

NOTES.—Data from Periods 1 to 8. Explanatory variables are at group level. Variable definitions: the construction of the Group cohesion index is explained in the paper (Section III.A); see SM1 for the Homophily Index and SM11 for Share. 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. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

In the three panels of Table SM11.1, we present four models exploring the impact of this share variable in the presence or absence of the other independent variables, which feature in the regressions of Table 2 in the main paper. In line with the analysis reported in Table 2 in the main paper, the Table-SM11.1 models are also estimated by pooling data for all 8 periods and using standard ordered probit with clustering at group-level (since groups make multiple decisions in the experiment).

The results provide some evidence that share matters, especially when it enters alone and 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 five out of six cases where both variables appear. The exception is Model 8 (Study-2 data only) which includes group cohesion, the homophily index and share. In this case both, group cohesion and share remain significant but only at the 10 percent level. The relatively weak significance here is plausibly a consequence of inter-correlations among the three regressors.[3] While these results are broadly 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 both Section VII in our paper and Section SM12 below, we underpin this conclusion more solidly, through an experimental manipulation which tests the sharing account more directly.

SM12: Bonus Study and Share Study: Additional Graphical Evidence

Here, we report graphical analysis compactly summarizing evidence from the two studies designed to benchmark group cohesion against two other variables: financial incentives ("Bonus Study"), and the likelihood of sharing payoffs ("Share Study"). We also provide further details of how those studies were conducted.

In the Bonus Study, we implemented four new, between-subjects, experimental treatments in which we varied the payoff matrix of the weak-link games by changing the bonus *b* of the payoff function (see Table SM3.1 above), while 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 fresh participants individually (not as groups of friends) who played 8 periods of the weak-link game, followed by a subset of the control tasks used in the Study-1 experiment (however, there was no elicitation of oneness ratings). In our four treatments, as shown in Table SM3.1, the bonus *b* was set

[3] The values of the pairwise correlation coefficients (pooled data) are, respectively: group cohesion vs homophily index, 0.610; group cohesion vs share (principal component), 0.441; homophily index vs share (principal component), 0.394. All three correlations are significant at $p < 0.001$.

equal to 6, 14, 22 and 30, respectively. A total of 240 new subjects participated in the Bonus Study (60 per treatment).

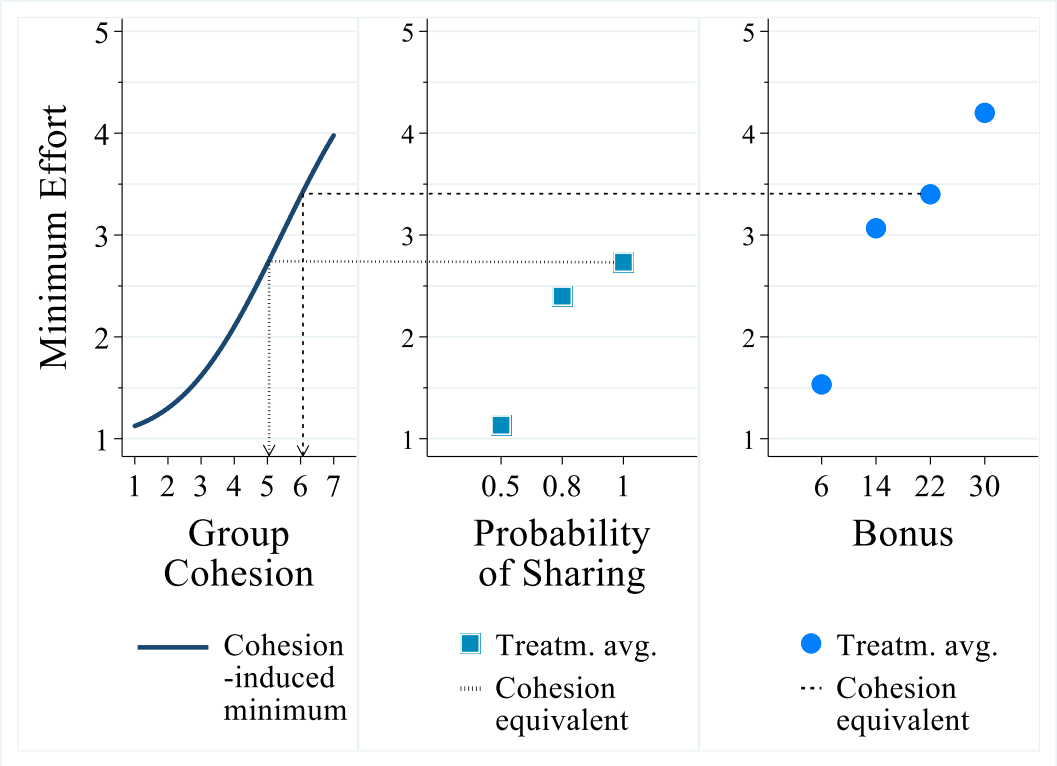


FIG. SM12.1.—The supply of minimum effort as a function of group cohesion in comparison with sharing rules and financial bonuses (group-level data from period 8). In the left panel, the minimum effort on the *y*-axis represents the expected minimum based on the predicted probabilities displayed in Fig. SM9.1c, using data from Studies 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 illustrates 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.

The right-most panel of Fig. SM12.1 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. SM12.1 where we plot the expected levels of minimum effort as a function of group cohesion based on the predicted probabilities displayed in Fig. SM9.1c (the "supply curve" of minimum effort as a function of group cohesion). The average minimum effort corresponding to a bonus of 6 is close to the lowest feasible effort (i.e., effort equal 1) and coincides with the expected minimum effort for low cohesion groups (i.e., group cohesion of approximately 3). Fig. SM12.1 below shows that sizable increases in the bonus are necessary to induce average minimum effort levels like those associated with high group cohesion. For instance, as Fig. SM12.1 illustrates (and as mentioned in our paper, too), a bonus of 22

yields an average minimum effort corresponding to that expected for group cohesion of approximately 6. This evidence reveals that the economic value of group cohesion is significant when benchmarked by the material incentives necessary to reach comparable average minimum efforts.

As a second benchmarking exercise, we ran the Share Study. As for the Bonus Study, fresh participants were recruited *individually* so that each subject typically did not know any other participant. In the Share Study, subjects completed 8 periods of the weak-link game with payoffs generated by setting the bonus *b* equal to 6 (see Table SM3.1). The distinctive feature of the Sharing Study was that, before playing the weak-link game, participants were informed that there was some probability that we would pool all individual earnings within each group and share them equally among group members. (All other standard procedures used across our studies were followed in the Share Study, too.) Part of the interest in running these treatments was hinted at in Section SM11 above 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 experimental treatments (60 subjects per treatment) in which we varied the known probability of sharing set at 0.5, 0.8 or 1, respectively. These treatments allow us to assess both an upper bound for the impact of sharing (when sharing is certain) and its sensitivity to different levels of uncertainty related to a potential sharing arrangement.

The results for these three treatments are presented as the middle panel of Fig. SM12.1. The treatment where the probability of sharing is equal to 1 (i.e., sharing is certain) produces an average minimum effort of 2.73, which is shown on the figure to correspond to the expected minimum effort associated with a group cohesion of approximately 5. When the probability of sharing is 50%, average minimum effort dropped dramatically to an expected minimum level only slightly above 1. Although this evidence on its own does not completely rule out the possibility that expectations of sharing might have played some role in our data, it weighs against it being a plausible account of the broad behavioral patterns in our data. This is partly because the ceiling of the sharing effect is substantially below the predicted effect of the highest cohesion level of 7 and because some degree of uncertainty – quite likely to be integral to any actual sharing arrangements – quickly reduces the impact of sharing.

SM13: Study-2 Mediation Analysis: Background and Complementary Analysis

In this section, we provide further discussion of the mediation analysis reported in Section VII: we expand on our approach and the motivation behind it; we also provide some additional analysis, complementing that in the main text.

While our empirical strategy for measuring group cohesion has built upon tools from social psychology, in seeking to understand the association between group cohesion and group minimum effort 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, 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 of the main paper 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, as shown in our paper, 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 social preferences.

The mediation analysis exploits individual-level data, collected as part of Study 2, on (i) beliefs about the effort decisions of other group members and (ii) their social preferences towards their group members. Each subject stated their best guess about what would be the minimum effort in a given period, straight after making their effort choice in that period but before receiving any relevant feedback. Then, towards the end of the study, we elicited participants' social preferences using the Social Value Orientation tasks: 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.[4] As mentioned in Section VII, we calculated

[4] 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 below, section SM14.c, for further details.

group-level beliefs as the group average of the individual beliefs per group and 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 using group-level data via the three linear equations presented in Section VII of the paper.

Inspired by the rational choice perspective, the model posits beliefs and (social) preferences as proximate determinants of minimum effort (as per Eq. 1 in Section VII). Beyond these standard influences, the model allows group cohesion to influence effort directly (as per Eq. 1) and indirectly via beliefs (Eq. 2 in Section VII) and social preferences (Eq. 3 in Section VII). The full logic of the model is most easily illustrated, graphically, via Fig. SM13.1. 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 association between group cohesion and effort.

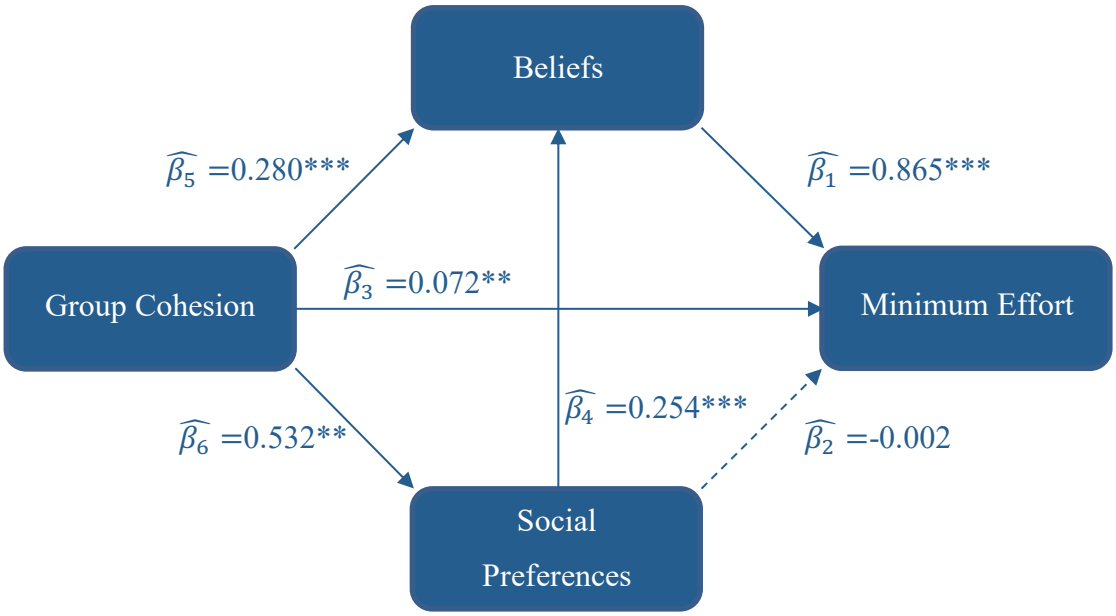


FIG. SM13.1.—Study-2 Mediation Analysis: modelling how group cohesion affects minimum effort. The figure reports econometric estimates of the simultaneous linear equations 1-3 above (all coefficients are standardized; with clustering at group level). Solid arrows are used where we detected statistically significant effects; a dashed arrow is used for the case where no statistically significant effect was detected.

Estimated parameters are included on the figure adjacent to their relevant branch. We find a statistically significant positive effect along every branch except for the one linking social preferences directly to minimum effort (this link is marked with a dashed arrow). Hence, as stated in the main text, 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.

TABLE SM13.1

MEDIATION ANALYSIS: ESTIMATING SIMULTANEOUS LINEAR EQUATIONS

Equation 1		
Dep. variable: Minimum Effort	(1)	(1')
Beliefs	0.865***	0.593***
	(0.025)	(0.098)
Social Preferences	-0.002	0.134
	(0.039)	(0.111)
Group Cohesion	0.072**	-0.033
	(0.033)	(0.122)
Equation 2		
Dep. variable: Beliefs	(2)	(2')
Social Preferences	0.254***	0.149
	(0.080)	(0.116)
Group Cohesion	0.280***	0.477***
	(0.100)	(0.103)
Equation 3		
Dep. variable: Social Preferences	(3)	(3')
Group Cohesion	0.532***	0.532***
	(0.068)	(0.080)
Log-likelihood	-4272.3	-536.3
# level 1 (resp. 2) units	552 (69)	69

NOTE.—Regressions 1-3: Study-2 data from Periods 1 to 8; standardized coefficients; clustering at group level; robust standard errors in parentheses. Regressions 1'-3': Study-2 data, period 1 only; standardized coefficients; standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$ * $p < 0.1$.

Table SM13.1 shows the estimated standardized coefficients (with standard errors) for the three-equation model. The first column of results (i.e., estimated equations 1, 2 and 3) uses data from periods 1 to 8 and the results correspond with those in Fig. SM13.1 above. Since beliefs in periods 2 to 8 are endogenous to game play, as a way of exploring possible dynamic effects, we re-estimated the mediation analysis which underpins Fig. 6 in our paper and Fig. SM13.1 but using only data from period 1. The corresponding estimates for this restricted sample are regressions 1' to 3' in SM13.1.

Comparing the estimates based on just period 1 to those using data from all the eight periods above, provides a window on the role played by dynamic effects. In the new estimates (1'-3' in Table SM13.1 above), although the Equation-1 coefficient on beliefs reduces from 0.865 to 0.593, the lion's share of the work is still done by beliefs and the coefficient remains highly statistically significant. The social preference coefficient in Equation 1 increases a little but fails to meet any conventional statistical significance threshold. The Equation-1 group cohesion coefficient falls and loses significance. However, the reduction in the direct effect of group cohesion on minimum effort as shown by Equation-1 estimates is counterbalanced by a substantial increase in the effect of group cohesion on beliefs (see Equation-2 estimates). In Equation 2, the social preference coefficient is substantially reduced and loses significance. The estimates for Equation 3 are essentially constant. As expected, given the large reduction in sample size, the log-likelihood across the two mediation analyses worsens for the estimates relying only on period 1 (it has only 69 observations compared with 552 for the full sample). So, to sum up, while there is some evidence of a dynamic effect, using only period 1 data does not substantially change the balance of explanatory power between beliefs, social preferences and group cohesion.

SM14: Experimental Materials and Instructions

The materials and instructions used were very similar for Study 1 and Study 2. Because of this we do not present the full set of materials separately for both studies. We do, however, explain where there were significant differences between them:

- Section SM14.a reports the invitation letter for the Study-2 experiment.
- Section SM14.b documents the wording used for the welcome of Study-2 experimental participants and the initial oral instructions.
- Section SM14.c presents the Study-2 experimental computerized instructions (note that only the key z-Tree screenshots are shown).
- Section SM14.d details the Study-2 post-experimental control tasks.

A difference between the studies was that in Study 1, participants' risk (à la Holt and Laury (2002)) and loss (à la Gächter, Johnson and Herrmann (2022)) attitudes were elicited while in Study 2, participants' game beliefs and social value orientations were elicited instead. Otherwise, Study-2 experimental procedures and instructions were identical to those used in Study 1 except for minor changes necessary to reflect location-specific details.

SM14.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

SM14.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?

SM14.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.

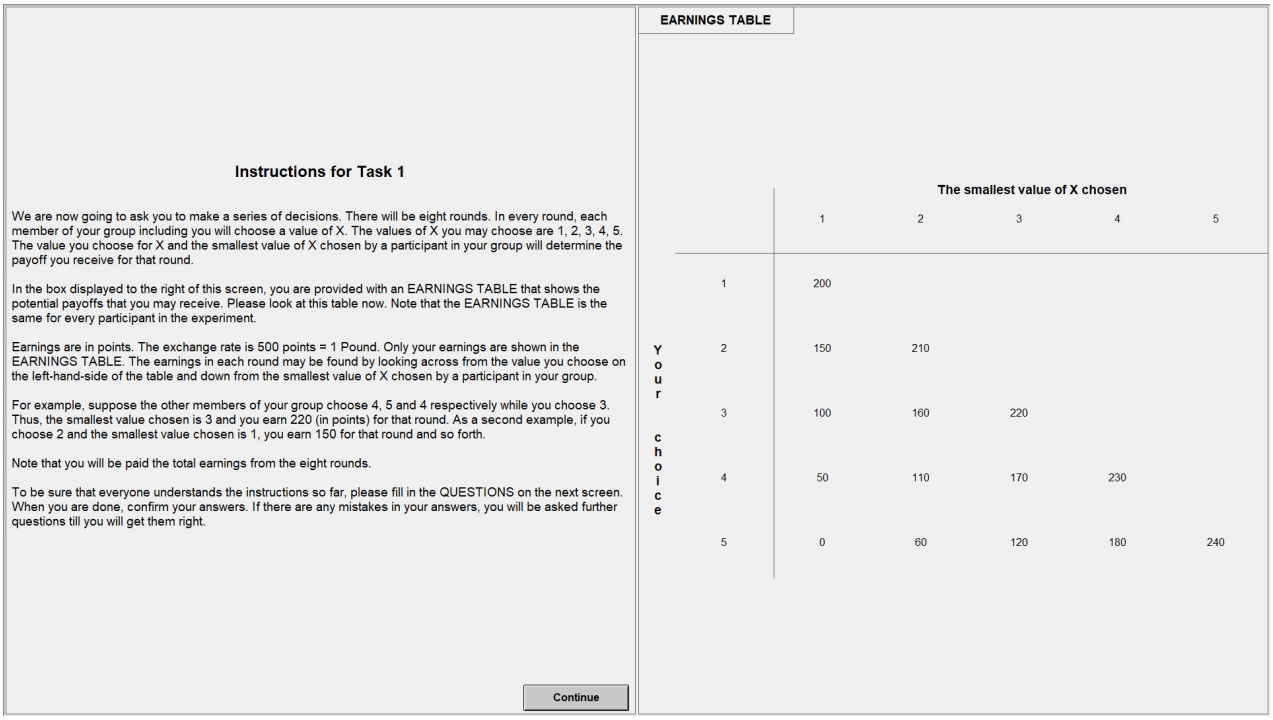
Continue [*Button*]

EARNINGS TABLE

		The smallest value of X chosen				
		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

****

Fig. SM14.1 below presents the screenshot of the "Instructions for Task 1."



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 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

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. SM14.1.—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				
		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

****

Fig. SM14.2 below is the screenshot of the input screen for the first period of the weak-link game.

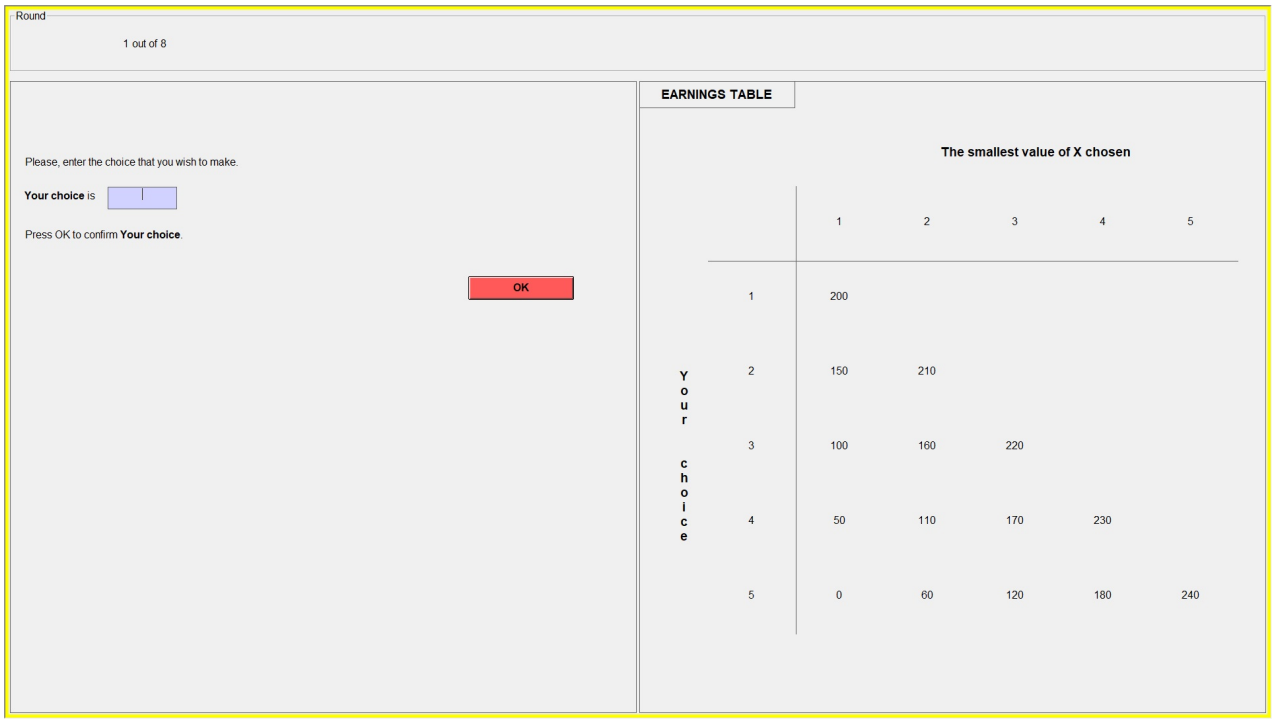


FIG. SM14.2.—Period 1 of the weak-link game: input screen.

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 ○ ○ ○ ○ ○ ○ ○ Very confident
Press the red button below to confirm your responses.
Confirm [*Button*]

EARNINGS TABLE

The smallest value of X chosen

		1	2	3	4	5
Your Ch	1	200				
	2	150	210			
	3	100	160	220		

o i c e	4	50	110	170	230	
	5	0	60	120	180	240

****

Fig. SM14.3 below reproduces the input screen for the first period belief and confidence elicitation.

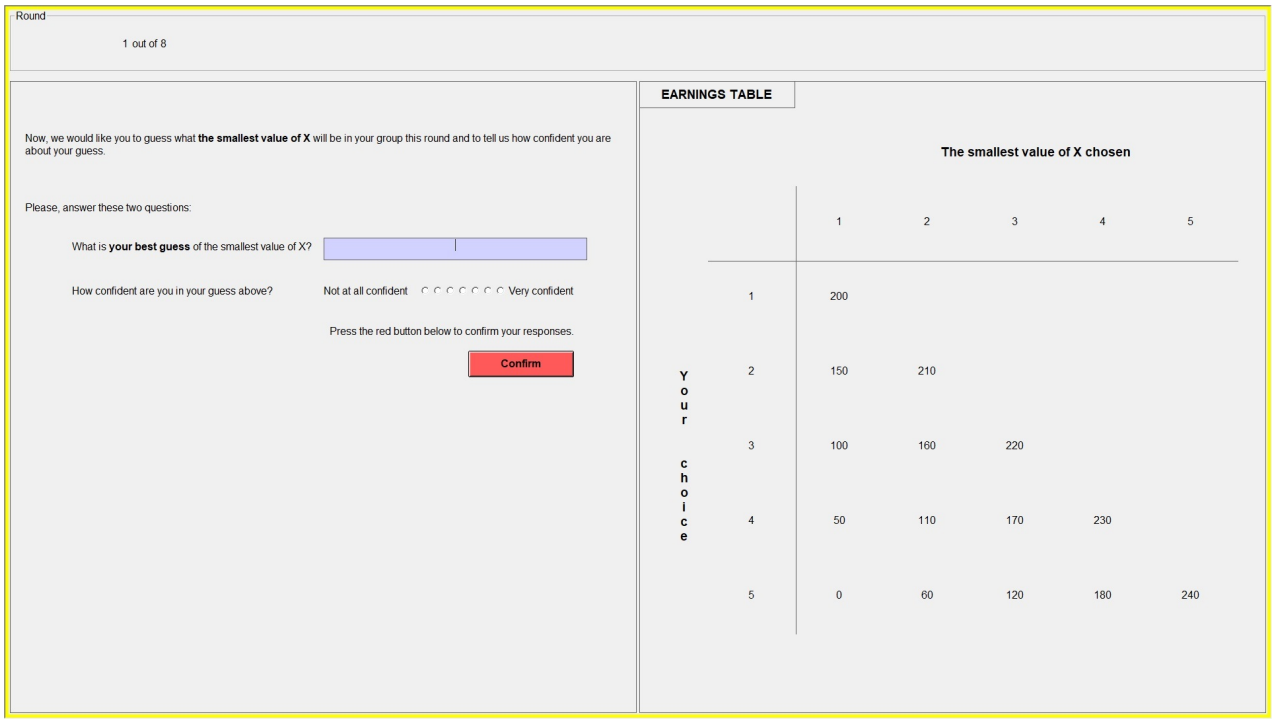


FIG. SM14.3.—Belief elicitation in period 1: input screen.

In each period, after each participant had entered their guess of the smallest value of X and stated their confidence, their computer screen reported 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 [*Output field*] 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.

	○	○	○	○	○	○	○	
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. SM14.4 below reproduces the screenshot eliciting the IOS rating.

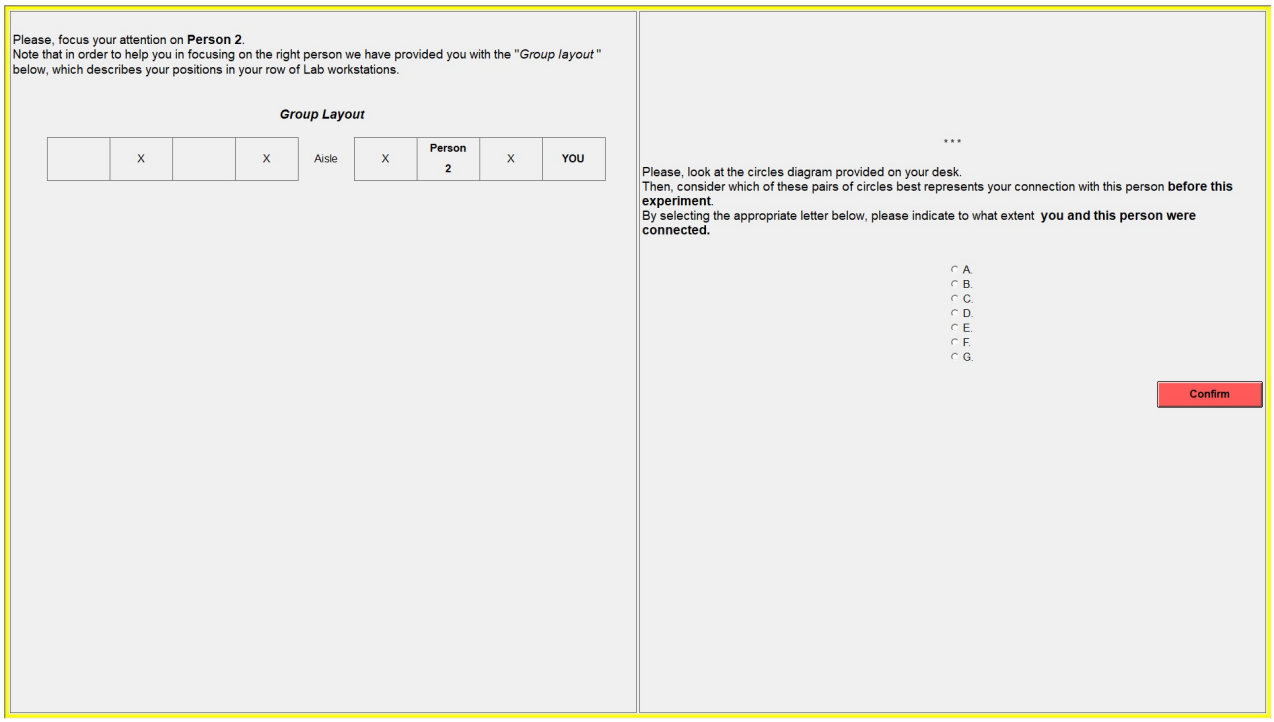


FIG. SM14.4.—IOS scale: input screen. Rater: Person 1; Focus Person: Person 2. The group was allocated to four workstations in a row having eight workstations of which only four were switched on. The positions of the switched-off workstations were marked by '×'.

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
	1	2	3	4	5	6	7

7. Warm	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
8. Concerned	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
9. Distressed	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
10. Low-spirited	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
11. Intrigued	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
12. Compassionate	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
13. Upset	o	o	o	o	o	o	o
	1	2	3	4	5	6	7
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
23. Bothered	o	o	o	o	o	o	o
24. Kind	o	o	o	o	o	o	o
25. Sad	o	o	o	o	o	o	o
26. Touched	o	o	o	o	o	o	o
27. Fortunate	o	o	o	o	o	o	o
28. Guilty	o	o	o	o	o	o	o
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, Weisel and Winter (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.

You receive	30	35	40	45	50	55	60	65	70	**You receive**	**50**
Other Person receives	80	70	60	50	40	30	20	10	0	**Other Person receives**	**40**

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. SM14.5 below reproduces the screenshot eliciting participants' risk attitudes.

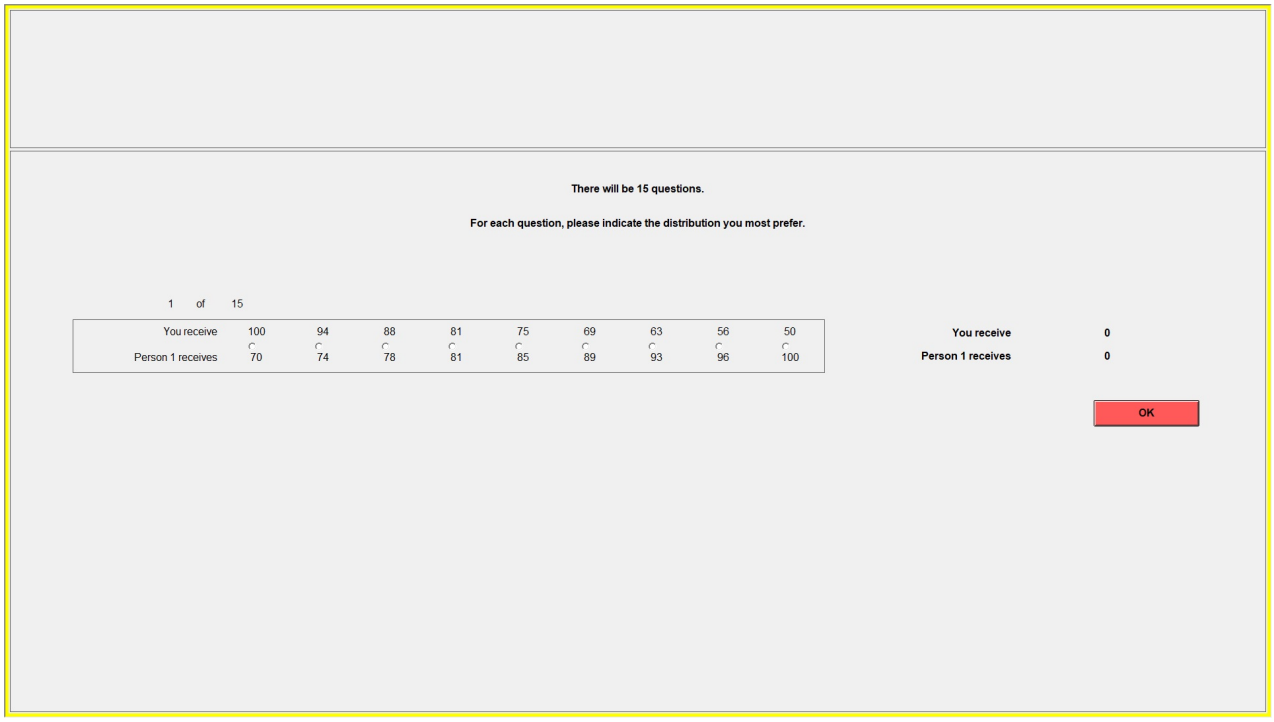


FIG. SM14.5.—SVO elicitation: question 1 out of 15 – input screen.

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.

SM14.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)

SM15: Experimental Procedures and Protocol

SM15.a. Inviting Participants

1. Roughly a week prior to the relevant experimental session, we send the invitation letter via ORSEE to several randomly selected participants, who have not taken part in the past in Coordination Games or in a similar experiment.
2. In what follows as well as in the correspondence with participants, we distinguish between signing-up, which is the act of enrolling in a session via ORSEE, and registration, which corresponds to communicating the full names of the three friends (as well as any other relevant detail, e.g., email addresses) after enrolling in a session.
3. We monitor participant registrations/sign-ups and, depending on the flow of registered/signed-up participants, we may send more invitations out.
4. In the correspondence with (potential) participants, we use letter templates as follows:
    a. The "Complete Your Registration" letter is sent to participants who have signed up via ORSEE, but have not yet communicated their names (and any additional detail required) to the experimenter;

b. The “Thanks a lot for communicating the full name…” letter is sent to participants who have completed their registration;
c. The “Registration Incomplete” letter is sent to those who have signed up via ORSEE but have not provided the experimenter with the full names of the three friends.

5. The recruitment of participants requires a continuous monitoring of both ORSEE and the experimenter’s mailbox.
6. It is also very important to constantly check for duplicates in the list of participants to avoid any returning subject. To do so it useful use the email alias of the participants (that can be checked for correctness in outlook and should uniquely identify each student).

SM15.b. Preparing the Economic Laboratory

The only preparation which needs to be undertaken beforehand is sticking with Sallotape a “Circles Diagram” on the desk of the workstation (preferably, on the left-hand-side in order to avoid the side where usually the computer mouse will be). To get a circles diagram, after printing the document “Circles Diagram”, cut along the dashed line on the printed document “Circles Diagram” and you will have two circles diagram copies to be stuck on two desks.

SM15.c. Preparing the Experimental Session

Open all the relevant z-Tree codes. Start the z-Leafs. Incidentally, note that it is important to sort the z-Leafs such that the participation number on the workstation corresponds to the subject number in z-Tree. Open the clients table.

SM15.d. Running the Session

Upon arrival participants are checked in against the Participant List and their ID is verified asking them to show their student ID card. Then, they are called forward (in order of arrival) in groups of friends. According to the treatment, they might be allocated in the same experimental group (if participating in the Friend matching treatment) or in four different groups (if participating in the Non-Friend matching treatment). The allocation of the workstations is at random by making participants pick a card (from a set of 4 cards each) with the workstation numbers. Of course, the set of 4 cards will differ according to the treatment. For instance, the first set of 4 cards for the Friend treatment (*F-treatment*, hereafter) might be 1, 2, 3, and 4; conversely, the first set of 4 cards for the Non-Friend treatment (*N-treatment*, hereafter) might be 1, 5, 9, and 13. It may be possible to run both treatment in

the same session: for instance, suppose you have 20 workstations, you may run the first 4 groups according to the N-treatment and the fifth group according to the F-treatment. (Indeed, the N-treatment can be run only with a number of groups that is divisible by four). It is important to keep a record of which group is participating in which treatment.

After participants have been sited, you can start the experiment by reading the oral instructions. Then, you should check that the number of subjects in the Background of the z-Tree codes is correct.

You can start the z-Tree codes in sequence beginning from Part 1 as far as Part 3. Then, you should run the questionnaire. In running Part 2, it may happen that a subject may have a question about the meaning of given terms describing emotions in the "Emotional Response Scale"; usually, you should have a browser open on an online dictionary on your mobile and read to the participant the dictionary definition.

Once participants have completed the questionnaire, they are asked to remain sited and quiet. We then call them, one at the time by their participation number. We ask them to come to the desk with all your belongings; at the desk, we pay them their experimental earnings in cash and in private. Then, participants are free to leave the lab.

SM References

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

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

Brandts, Jordi, and David J. Cooper, "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 (2007), 1223-68.

Brandts, Jordi, David J. Cooper, and Enrique Fatas, "Leadership and Overcoming Coordination Failure with Asymmetric Costs," *Experimental Economics*, 10 (2007), 269-84.

Brandts, Jordi, David J. Cooper, and Roberto A. Weber, "Legitimacy, Communication, and Leadership in the Turnaround Game," *Management Science*, 61 (2015), 2627-45.

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

Crosetto, Paolo, Ori Weisel, and Fabian Winter, "A Flexible Z-Tree Implementation of the Social Value Orientation Slider Measure (Murphy Et Al. 2011) – Manual," Jena Economic Research Paper # 2012-062. Available at SSRN: http://dx.doi.org/10.2139/ssrn.2195598 (2012).

Gächter, Simon, Eric J. Johnson, and Andreas Herrmann, "Individual-Level Loss Aversion in Riskless and Risky Choices," *Theory and Decision*, 92 (2022), 599-624.

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

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

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

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

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

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