Center for European, Governance and Economic Development Research
cege

Discussion Papers

Number 185 – January 2014

THE IMPACT OF COMMUNICATION REGIMES ON GROUP RATIONALITY: EXPERIMENTAL EVIDENCE

Lukas Meub, Till Proeger

Georg-August-Universität Göttingen

ISSN: 1439-2305

THE IMPACT OF COMMUNICATION REGIMES ON GROUP RATIONALITY:

EXPERIMENTAL EVIDENCE

Lukas Meub[1], **Till Proeger**

Faculty of Economic Sciences, Chair of Economic Policy and SME Research, University of Göttingen

Abstract: The performance of groups has been thoroughly investigated in experimental economics, showing that groups are overall more rational deciders than individuals. However, superior group performance in economic experiments has primarily been shown for face-to-face decision making, which has ceased to be the prevalent form of communication in many IT-based organizations. To test the robustness of higher group rationality under conditions of virtual communication, we conduct a social learning experiment. We find that virtual communication leads to a substantial deterioration of group rationality for a judgmental task, while there is no effect for a purely intellective task. Further, we show that higher cognitive abilities of group members have no impact for the judgmental task, yet increase rationality for the intellective task. Our results have potential implications for the design of communication structures within decentralized organizations relying on virtual communication.

Keywords: cognitive abilities; communication; group composition; group performance; laboratory experiment; methodology; social learning

JEL Classification: C9; D8

[1] Corresponding author: Lukas Meub, Faculty of Economic Sciences, Chair of Economic Policy and SME Research, University of Goettingen, Platz der Goettinger Sieben 3, 37073, Goettingen, Germany, e-mail: lukas.meub@wiwi.uni-goettingen.de, phone: +49 551 39 7761; fax: :+49 551 39 19558.

1. INTRODUCTION

Experimental economics' understanding of group performance has advanced significantly in the past decade. Comparing individual and group results for the major experiments in behavioral economics, the core finding is that groups are more rational deciders overall (Kugler et al., 2012). By overcoming cognitive limitations, groups are shown to be "less behavioral *than individuals"* (Charness and Sutter, 2012, p. 159) and thus regularly closer to theoretically rational solutions. This is interpreted as a partial vindication of the assumptions on rational behavior for the prediction of real-world economic decision making.

By incorporating group cooperation in economic contexts, behavioral research more accurately reflects the widespread team decision making in actual organizations. However, economic experiments to date have given little consideration to the effects of different external restraints on group decision making. Specifically, as pointed out in Kugler et al.'s (2012) literature review, the impact of varying modes of communication is yet to be investigated, particularly for the case of virtual communication and personal group discussions. This specific issue has gained relevance with the fundamental shift in the nature of organizational decision making towards digital communication and decentralized decision making. An increasing share of team decisions in international organizations is taken in geographically dispersed locations, using means of virtual communication rather than personal meetings (Balliet, 2010). To date, economic experiments have not considered this aspect's potential consequences for group rationality.[2] Of the thirty-seven economic group experiments reviewed in Kugler et al. (2012), thirty implement direct face-to-face discussions or a no-communication voting procedure. While seven studies feature computer-mediated discussions, none focusses on potential differences from face-to-face interaction.

In contrast, empirical studies in organizational psychology and field studies on IT-management have covered the effects of virtual decision making, overall pointing to rather negative consequences for team performance, e.g. in the meta-study by Baltes et al. (2002). De Guinea et al.'s (2012) recent contribution points to overall negative effects, particularly in short-term teams. Mesmer-Magnus et al. (2011), using the extent of information sharing as a measure for group success, show that low levels of virtuality are supportive, while increased

[2] In contrast, group cooperation in social dilemma situations with computer-mediated communication has been investigated, e.g. by Bicchieri and Lev-On (2007). Accordingly, the meta-study by Balliet (2010) finds face-to-face communication to be more effective in fostering group cooperation than virtual communication.

levels rather constrain effective information sharing. However, the derivation of clear behavioral patterns remains difficult, due to the ubiquitous lack of clear theoretical benchmarks and financial incentives (Charness and Sutter, 2012). The same holds true for the broad results of small group research in social psychology (for a review see Laughlin, 2011), which suggests that groups are more successful if the demonstrability of the respective correct solutions is high (Laughlin et al., 2002). The performance of groups can thus be predicted by the respective task's position on a continuum from intellective to purely judgmental (Laughlin, 1980).[3] With demonstrability as the core influence on group performance, virtuality might reduce a group's ability to effectively convey information, leading to the deterioration of performance and thus reduce group rationality premiums.

We build on this notion to further economic small group research by investigating the robustness of higher group rationality when solutions have to be agreed upon by means of virtual communication. Therefore, we replicate a recent study on group performance in a social learning game that features personal communication presented by Fahr and Irlenbusch (2011, henceforth: F&I). All other factors constant, we have participants communicating and deciding via chat in two treatments, one featuring a purely intellective task and another that introduces a judgmental element. By contrasting both studies, we assess the impact of virtual communication on group rationality. Our secondary aim in this paper is to investigate the impact of group composition according to cognitive abilities on group rationality. Such group composition effects have not been considered to date in economic group research (Kugler et al., 2012), yet seem useful to give insight into reasons for group success. This might help to determine whether groups in economic contexts tend to perform like their most capable members or if the process of group discussions itself is pivotal for superior group performance by generating additional insight.

In comparison to the benchmark data from F&I, we find virtual communication to have a profound impact on decision quality. Virtual communication considerably reduces the rationality of group decisions for a judgmental task, while there is no effect for the purely intellective task. Secondly, higher cognitive abilities of group members have a positive effect for the intellective task and no effect for the judgmental task. Accordingly, even high cognitive abilities in groups fail to eliminate the detrimental effect of virtual communication in the judgmental task. Overall, in the case of judgmental tasks, we can conclude that groups

[3] Intellective tasks possess a clear normative criterion to evaluate the quality of subjects' decisions, while judgmental tasks involve uncertainty about the actual quality of decisions (Cox and Hayne, 2006).

rely on face-to-face group discussions to achieve superior rationality when compared to individuals. For a purely intellective task, superior group performance holds regardless of the communication regime.

The remainder of this paper is organized as follows. In section 2, the experimental design is introduced, before section 3 presents our behavioral hypotheses. Section 4 describes our results and section 5 draws a conclusion.

2. EXPERIMENTAL DESIGN

Our experimental design replicates F&I, who built their group study on the urn game by Anderson and Holt (1997) and the respective seminal herding models by Banerjee (1992) and Bikhchandani et al. (1992). Using this experimental setting as a working horse yields the advantage of building on a large and well-established strand of literature on social learning, with Weizsäcker (2010) offering a comprehensive meta-study. As in the seminal game, two urns are considered, one containing two red and one blue marbles and the other two blue and one red. One urn is randomly selected prior to each period. Six subjects sequentially and in a randomized order guess which urn was selected. Along the decision sequence, all guesses become public information. Before taking a choice, each player is shown one marble, which is randomly drawn from the selected urn. At the end of every period, the correct urn is revealed. Rational decisions require correct Bayesian updating. We base our analysis on the basic benchmark model of rational behavior employed by the studies following Anderson and Holt (1997).[4]

There are two treatments: in the private information treatment (PIT), players observe all prior decisions taken in the sequence, as described above; in the full information treatment (FIT), players are instead shown the marbles drawn by previous players. As deducted by F&I, this setting allows for a distinction between judgmental and intellective tasks. In FIT, subjects have to correctly apply a counting rule regarding prior drawn marbles, a considerably straightforward intellective task. By contrast, in PIT, subjects additionally need to interpret the observed decisions with respect to the preceding players' decision quality. Thus, there is an additional element of uncertainty, which serves as a judgmental aspect to the otherwise intellective task.

[4] See e.g. Goeree et al. (2007) for a comprehensive theoretical analysis.

In addition to the six individual players, a group of three players is matched to one of the six individual players ("focal individual") during the entire game. Both the group and the focal individual face identical situations in all periods. Unlike the focal individual's the group decision does not become public information. In each of the fifteen periods, groups are required to come up with a unitary solution. In both treatments, guessing the correct urn pays 1€ for each group member.

To closely replicate F&I, we implement the decision sequences generated in their treatments. We match groups of three players and have them decide in the same situations as their groups. It was common knowledge that the decision sequences shown to groups were generated by individual players facing the identical game. This procedure provides us with a large number of observations for group decisions in similar situations. While all other factors remained constant, we had groups communicate via chat rather than deciding face-to-face, as in the benchmark experiment. Henceforth, the benchmark treatments are denoted as FIT/PIT f2f, in contrast to FIT/PIT chat. To rule out the possibility that distrust in the preprogrammed decision sequences drive our results, we ran an additional treatment reported in subsection 4.3.

Furthermore, as we aimed to determine the impact of varying cognitive abilities within groups, participants took the cognitive reflection test (Frederick, 2005)[5] beforehand, whereby subjects were paid 1€ for answering all three questions correctly. Groups were subsequently matched according to their performance in the test. We classify individuals in two categories, as in Oechssler et al. (2009): participants with two or more correct answers in the CR-test and those with one or less correct answer. Groups exclusively composed of participants of the first category are referred to as high, groups composed of the second category as low and groups composed of both categories as mixed.

For both treatments there were four sessions, giving us a total of 93 (24/21/24/24) participants in PIT and 90 (24/24/21/21) for FIT. For the additional treatment serving as a robustness check, we conducted 8 sessions, with 9 subjects each. We used zTree (Fischbacher, 2007) and ORSEE (Greiner, 2004). Experiments were carried out at the Göttingen Laboratory of Behavioral Economics using a standard subject pool across all disciplines. 46% of participants

[5] The cognitive reflection test by Frederick (2005) is a simple three-item measure testing the individual tendency of choosing intuitive responses over more reflected solutions. The CRT is substantially correlated with cognitive abilities and also a good predicator of performance on heuristics-and-biases tasks, as reported by Toplak et al. (2011).

in our sample were female. Including show-up fees, subjects earned 10.40€ on average. The sessions lasted around sixty minutes.[6]

3. HYPOTHESES

Given that group success crucially depends on the demonstrability of correct solutions (Laughlin et al., 2002), our basic hypothesis is that the mode of communication has a detrimental influence on overall rationality when compared to the baseline face-to-face setting by F&I. This is expected to hold for both treatments, i.e. the purely intellective and the judgmental task. We thus assume that the positive effects of group decision making will not hold regardless of the mode of communication for a social learning game. Therefore, we formulate:

Hypothesis 1 ("Rationality and virtual communication").

a) Virtual communication reduces rationality in the full information treatment (FIT).
b) Virtual communication reduces rationality in the private information treatment (PIT).

Our second hypothesis is concerned with the effect of group composition for the particular case of cognitive abilities. Since higher cognitive abilities are associated with more rational behavior on an individual level (see e.g. for the beauty contest game, Burnham et al., 2009; Brañas-Garza et al., 2012), we expect groups composed of subjects scoring well on the CR-test to achieve more rational decisions. We thus imply that it is not the process of group discussion itself, but rather the individual cognitive abilities that increase group performance.

Hypothesis 2: ("Rationality and cognitive abilities").

a) Groups composed of individuals showing higher cognitive abilities act more rationally in the full information treatment (FIT).
b) Groups composed of individuals showing higher cognitive abilities act more rationally in the private information treatment (PIT).

[6] The original instructions were in German and are available from the authors upon request. A translation is documented in the appendix.

4. RESULTS

For the basic analysis of rationality in group decisions, we build on the same assumptions as F&I. Rational behavior requires players in PIT to identify and correctly interpret informative and uninformative prior decisions. A decision is uninformative, i.e. reveals no additional information, if Bayesian updating prompts a preceding player to pick a specific urn regardless of their private information. Rational players thus apply a counting rule restricted to informative decisions.

To qualify decisions as being in line with Bayesian rationality, it is assumed that all players generally believed that their predecessors decided in accordance with Bayes' Rule. The distinct violation of this behavior in a PIT setting, i.e. a decision not following a cascade of prior decisions, is assumed to prompt subsequent deciders to believe that respective individuals followed their private signal.[7] Also, if the restricted counting rule leads to an indifference situation regarding which urn to choose, players are assumed to expect previous deciders in these situations followed their private signals.[8]

By contrast, in FIT, rational decisions imply the use of a simple counting rule, as all prior draws are informative. Again, if applying the counting rule gives the same probability for both urns, players are expected to decide in accordance with their private signals.

Both treatments include situations whereby the private signal and the majority of preceding drawn marbles or decisions are congruent, as well as situations in which the private signal opposes the decision history. We denominate a situation whereby players need to discard their private signals in order to act rationally as a non-private situation. Past studies have pointed out the relevance of these decision situations, with the reluctance to discard private signals being the main cause for irrational decisions in the urn game (Weizsäcker, 2010). We thus put an emphasis on the analysis of the varying performance in non-private situations.

With regard to the CR-test taken by participants beforehand, we find that 25.68% of all participants answered none of the three questions correctly, 19.67% give one and 27.87% two correct answers, while 26.78% achieve three correct answers. Accordingly, our matching

---

[7] A cascade is defined as a situation in which it becomes rational to follow the preceding pattern of decisions regardless of private signals. For instance, the player on the third position in a decision sequence rationally follows a uniform decision of the two preceding players regardless of her own private signal. Thus, a cascade situation is established, which continues for the rest of the period, rationally ending in a unanimous decision of all players.

[8] These assumptions are in line with behavior regularly observed in social learning games and are even more plausible when taking into account a small, but positive error rate (see e.g. Weizsäcker, 2010).

procedure gives us 37.7% high groups, 32.79% low groups and 29.51% mixed groups. Recall that all group types were presented similar preprogrammed decision situations, thus differences in rationality can be ascribed to differences in cognitive abilities.[9]

Since we are interested in comparing our data to the results obtained by observing randomly matched groups under face-to-face communication, we estimate a probability weighted average of decision rationality for our matched groups. This calculation follows a simple procedure accounting for the probability that a specific group type is formed under a random matching routine and the fractions of correct answers in the CR-test.[10]

Table 1 presents the aggregate data for decision rationality in our treatments and Table 2 summarizes the results of F&I. If not mentioned otherwise, all testing procedures are performed "two-sided" and every individual or group is treated as one observation only.

Treatment	**share rational (obs.)**				**share rational (obs.) in non-private situations**			
group type	**high**	**low**	**mixed**	**w-average**	**high**	**low**	**mixed**	**w-average**
FIT chat	.9758 (165)	.9267 (150)	.9556 (135)	.9650 (450)	1.0000 (25)	.8462 (26)	.9048 (21)	.9021 (72)
PIT chat	.9056 (180)	.8733 (150)	.9111 (135)	.8984	.5357 (28)	.3043 (23)	.5455 (22)	.5304 (73)

Table 1: Rationality by group type for both treatments with the total number of observations in parentheses. Percentages are given relative to the respective total number of observations. w-average denotes the weighted average.

[9] For FIT chat and PIT chat there are potentially 180 preprogrammed decision situations derived from the 12 groups of F&I. Each group is presented 15 situations. Thus, not all decisions are covered by each group type, e.g. for the mixed groups in PIT only 129 situations were actually played (149 for high groups, 141 for low groups). Differences between the actual number of situations played and the total number of observations occur due to a small number of situations that were played more than once, which was caused by programming restrictions of z-Tree. However, since there are no systematic differences in the situations presented to the different group types, we can rule out that varying color histories triggered differences in rationality on the group type level.

[10] The probability of a subject to answer one or less questions in the CR-test correctly amounts to 0.4535, consequently the probability of answering two or more correctly is equal to 0.5465. Thus, a high group should appear 16.32% (=0.5465^3) of the times given a random matching procedure; a low group in 9.33% (=0.4535^3) and a mixed group in 74.35% (=$1-0.5465^3-0.4535^3$) of the times. The weighting factors for high, low, mixed are then .4329, .2844, 2.52 respectively.

Treatment	share rational (obs.)		share rational (obs.) in non-private situations	
	groups	individuals	Groups	individuals
FIT f2f	.9278 (180)	.9093 (1080)	.7419 (31)	.707 (215)
PIT f2f	.9722 (180)	.9028 (1080)	.8966 (29)	.6436 (202)

Table 2: Rationality for groups and individuals as reported in Fahr and Irlenbusch (2011) with the total number of observations in parentheses. Percentages are given relative to the respective total number of observations.

4.1 PERFORMANCE IN FIT

Analyzing decision rationality with respect to the mode of communication, we find that the overall weighted average in FIT chat is not significantly lower in comparison to FIT f2f (Satterthwaite's unequal-variance t-test, t=-1.0708, p=.2927). In contrast, for non-private situations, group rationality is around 16 percentage points higher in FIT chat (Satterthwaite's unequal-variance t-test, t=-1.7932, p=.0952), pointing to a positive effect of virtual communication in non-private situations. Thus, the implementation of virtual communication might even foster rational decision making for a purely intellective task. However, we conclude that virtual communication does not lead to a deterioration of rationality for the full information condition (FIT) and thus find no evidence in support of H1a.

Considering differences with respect to cognitive abilities, we cannot reject the null of equal overall rationality across group types (Kruskal-Wallis test with ties, chi2(2)=3.421, p=.1807; for non-private situations chi2(2)=3.234, p=.1985). However, there is a significant difference in rationality between high and low groups (Mann-Whitney test, z=1.811, p=.0701; for non-private situations z=1.824, p=.0682). Mixed groups show a level of rationality, that is neither significantly different from low groups (Mann-Whitney test, z=-.0370, p=.7116) nor from high groups (Mann-Whitney test, z=.673, p=.5011). We conclude that there is, albeit weak, evidence in support of H2a since at least high groups outperform lows. High cognitive ability of all group members has a small positive impact on group rationality compared to groups in which all members show low cognitive ability.

4.2 PERFORMANCE IN PIT

In PIT chat, as indicated by the weighted average, group rationality is significantly lower compared to PIT f2f (Satterthwaite's unequal-variance t-test, t=3.3085, p=.0027). Additionally, groups in PIT chat show no significantly different overall rationality compared to individual players in PIT f2f (Satterthwaite's unequal-variance t-test, t=.3613, p=.7189). In non-private situations groups not only perform worse under virtual communication (Satterthwaite's unequal-variance t-test, t=4.0059, p=.0003), they show even lower levels of rationality when compared to individual players (Satterthwaite's unequal-variance t-test, t=1.7771, p=.0831). The rationality premium for groups communicating face-to-face amounts to 36.8 percentage points in non-private situations. In sum, we find strong evidence in support of H1b. Accordingly, groups using chat more often fail to discard their private signal in the crucial non-private situations, which leads to deteriorating overall rationality for the judgmental task (PIT).

Concerning cognitive abilities in PIT chat, we find that overall rationality is not significantly different across group types (Kruskal-Wallis test with ties, chi2(2)=1.046, p=.5929; for non-private situations chi2(2)=1.826, p=.4014). Differences between mixed and high groups are particularly small. In non-private situations, low groups discard their signals less often compared to high and mixed groups. This could be interpreted as supporting the notion that the most capable team members drive overall group rationality. However, the difference between pooled mixed and high groups on the one side and low groups on the other is not significant (Mann-Whitney test, z=.983, p=.3258; for non-private situations z=1.285, p=.1989).

Nevertheless, the rationality of high groups in PIT chat is still significantly lower compared to the randomly matched groups in PIT f2f (Mann-Whitney test, z=-1.780, p=.0751; for non-private situations z=-2.102, p=.0355).[11] Thus, even groups showing higher cognitive abilities are unable to compensate the negative effect of virtual communication. We conclude that there is no evidence in support of H2b, thus groups showing higher cognitive abilities are not capable of deciding more rationally in the judgmental task.

4.3 ROBUSTNESS CHECK FOR PIT

Obviously, it could be argued that subjects might have primarily suspected the preprogrammed decision histories in PIT chat to be manipulated. This distrust could in turn

[11] This result can be seen as giving additional support to the robustness of our main findings. Since subjects in the cognitive reflection test performed rather poorly in comparison to the results presented by Oechssler et al. (2009) and Frederick (2005), the estimated weighted averages may be somewhat too low. However, our main treatment effect of decreasing rationality due to the virtual communication regime proofs robust, since even the high cognitive ability groups are not capable to achieve similar levels of rationality as randomly matched groups communicating face-to-face.

have driven the reluctance to discard private signals, implying that virtual communication per se cannot be held responsible for the strong in rationality. To clarify this, we test for the robustness of our results by running an additional treatment that concisely replicates the benchmark treatment by F&I, but has groups communicate via chat. Consequently, groups observed decisions taken by individual players who were now physically present in the lab and groups are assigned to an individual player (focal individual), facing identical decision situations. We ran 8 sessions with 9 subjects, giving us 8 randomly matched groups each taking 15 decisions. We find 89.17% of group decisions to be rational with only 60.9% in non-private situations. The respective focal individuals achieve 96.66% rational decisions overall with 95.65% in non-private situations (Wilcoxon matched-pairs signed-ranks test, z=-1.616, p=.1060; for non-private situations z=-2.188, p=.0287).[12] Thus, focal individuals performed even better than the respective groups and not vice versa. At all, individuals show an overall rationality of 91.11% with 69.66% in non-private situations, which is not significantly different from the respective group performance (Mann-Whitney test, z=.788, p=.4307; for non-private situations z=.207, p=.836).

At the same time, group rationality is not significantly different from the groups in PIT chat (Satterthwaite's unequal-variance t-test, t=.4520, p=.6614; for non-private situations t=-1.5284, p=.1424). Nonetheless, it is again significantly lower than in PIT f2f (Mann-Whitney test, z=2.293, p=.0219; for non-private situations z=2.039, p=.0415).

We can thus rule out distrust in the preprogrammed decision histories as being the driving force for our treatment effect. Groups in PIT rely on personal communication to achieve relatively higher levels of rationality when compared to individuals.

5. CONCLUSION

Economic group research has arrived at the conclusion that groups are more rational economic agents, which may rehabilitate assumptions of rational behavior in actual decision situations. We argue that the experimental studies showing higher rationality need to be extended to encompass virtual communication, which has emerged as a prevalent form of

[12] As pointed out by F&I, tests within the PIT treatment regarding differences between individuals and groups have to be performed using a testing procedure for matched pairs (Wilcoxon signed rank test). This is due to the interdependence of groups' and the focal individuals' decisions. Recall that the focal individuals' decisions become public information and thus influence other individuals' decisions that are presented to groups as well as to focal individuals in subsequent periods. Accordingly, observations of groups and focal individuals within PIT are not independent.

decision making within organizations. To further this point, we compare the rationality of groups in a social learning game featuring face-to-face decision making and chat communication. We show that group rationality remains high for an intellective task, but strongly deteriorates for a judgmental task. Our group matching procedure shows that higher cognitive abilities lead to increased rationality for an intellective task but not for a judgmental task. Apparently, groups are able to perform better given more cognitively capable members in intellective tasks, regardless of virtual communication. For the judgmental task, in contrast, even high cognitive ability groups using virtual communication are unable to achieve the rationality of those using face-to-face communication. Thus, the process of group discussions appears to be critical for group rationality instead of the members' cognitive abilities. These results stress the importance of the demonstrability of solutions for superior group performance. Once judgmental reasoning is hindered by non-personal communication, group performance deteriorates even below the average level of individuals, which is in line with experimental results in social psychology (Laughlin et al., 2002). We would thus argue that superior group rationality in many economic studies involving judgmental aspects might not hold for conditions of virtual communication. Considering real-world decision making, our results hold relevance when choosing between different potential communication regimes for group decisions. Our results suggest that virtual communication is a suitable tool for rather simple decisions that have clear and demonstrable solutions. In this case, more capable members efficiently convey information, which enables groups to take better decisions. For judgmental tasks, personal conversations appear to foster more successful group decision making. While personal meetings may be more costly, we argue that they are bound to increase group performance for judgmental tasks. Inversely, if the organizational structure only allows for virtual communication, having individuals take decisions involving judgmental aspects might yield superior results. We thus argue that communication structures in decentralized organizations should be carefully crafted according to the specific tasks at hand in order to optimize group and individual performance.

APPENDIX

The instructions documented below refer to the PIT treatment. The differences to the FIT treatment are indicated in brackets.

General Information

In this game, there are two urns, each with three marbles. Urn A contains two red and one blue marbles. Urn B contains two blue and one red marbles. The number and distribution of the marbles will remain constant during the entire game.

At the beginning of every round, the computer will select one of the two urns (A or B), with equal likelihood. You are not informed which urn has been chosen for this round. Six players now have to guess sequentially which urn has been chosen. Once it is your turn, one of the three marbles will be chosen by the computer, with equal likelihood, and shown to you. Every player can only see the marble that has been drawn for him/her.

Besides the individual marbles, each player can observe [the previous marbles drawn for] the decisions of the other players. For instance, this means that the second player sees [the marbles drawn for the first player] the decision of the first player, the third player sees the [marbles] decisions of the first and second players, and so forth.

The course of the game

In this game, you do not decide on your own, but rather in a group with two other players. You can communicate using the chat function to take a unanimous decision. The chat and all other data will be recorded anonymously. At first, you will answer three questions, and will have two minutes to answer each question. Your group membership will be determined according to the results of the questions. Accordingly, the game will start once all players have answered the questions. With your group, you will be randomly assigned to one of the six positions in the decision sequence. The decisions [marbles drawn] you see are from a prior experiment with individual players, who played the game under the exact same conditions as you. After your group has been shown the randomly determined marble, you have to take a common guess within 3 minutes. Once all groups have taken their guesses, the round ends and the correct urn will be shown to all players of your group. This is repeated for 15 rounds.

Your payoff

For a correct guess, each group member receives a payoff of 1€. The gains from all rounds will be added and paid to you at the end of the game. You will additionally be paid 1€ for the correct answer to all three questions at the beginning. Please note: If your group fails to give a common answer within 3 minutes (and an additional 30 seconds), you will not receive a payoff for the respective round.

REFERENCES

Anderson, L.R., Holt, C.A., 1997. Information Cascades in the laboratory. American Economic Review 87 (5), 847-862.

Balliet, D., 2010. Communication and cooperation in social dilemmas: A meta-analytic review. Journal of Conflict Resolution, 54(1), 39-57. doi: 10.1177/0022002709352443.

Baltes, B., Dickson, M., Sherman, M., Bauer, C., LaGanke, J., 2002. Computer-Mediated Communication and Group Decision Making: A Meta-Analysis. Organizational Behavior and Human Decision Processes 87 (1), 156-179. doi: 10.1006/obhd.2001.2961.

Banerjee, A.V., 1992. A simple model of herd behavior. Quarterly Journal of Economics, 107 (3), 797-817. doi: 10.2307/2118364.

Bicchieri, C., Lev-On, A., 2007. Computer-mediated communication and cooperation in social dilemmas: an experimental analysis. politics, philosophy & economics, 6(2), 139-168. doi: 10.1177/1470594X07077267.

Bikhchandani, S., Hirshleifer, D., Welch, I. (1992). A theory of fads, fashion, custom, and cultural change in informational cascades. Journal of Political Economy 100(5), 992-1026. doi: 10.1086/261849.

Brañas-Garza, P., Garcia-Muñoz, T., González, R. H., 2012. Cognitive effort in the beauty contest game. Journal of Economic Behavior & Organization, 83(2), 254-260. doi: 10.1016/j.jebo.2012.05.018.

Burnham, T. C., Cesarini, D., Johannesson, M., Lichtenstein, P., Wallace, B., 2009. Higher cognitive ability is associated with lower entries in a< i> p</i>-beauty contest. Journal of Economic Behavior & Organization, 72(1), 171-175. doi: 10.1016/j.jebo.2009.05.015.

Charness, G., Sutter, M., 2012. Groups make better Self-Interested Decisions. Journal of Economic Perspectives 26 (3), 157-176. doi: 10.1257/jep.26.3.157.

Cox, J., Hayne, S., 2006. Barking up the right tree: are small groups rational agents?. Experimental Economics 9, 209-222. doi: 10.1007/s10683-006-9123-3.

De Guinea, A.O., Webster, J., Staples, D., 2012. A meta-analysis of the consequences of virtualness on team functioning. Information & Management 49, 301-308. doi: 10.1016/j.im.2012.08.003.

Fahr, R., Irlenbusch, B., 2011. Who follows the crowd - Groups or individuals?. Journal of Economic Behavior & Organization 80 (1), 200-209. doi:10.1016/j.jebo.2011.03.007.

Fischbacher, U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10, 171-178. doi: 10.1007/s10683-006-9159-4.

Frederick, S., 2005. Cognitive reflection and decision making. The Journal of Economic Perspectives 19 (4), 25-42. doi: 10.1257/089533005775196732.

Goeree, J.K., Palfrey, T.R., Rogers, B.W., McKelvey, R.D., 2007. Self-correcting information cascades. Review of Economic Studies 74 (3), 733-762. doi: 10.1111/j.1467-937X.2007.00438.x.

Greiner, B., 2004. An online recruiting system for economic experiments. In: Kremer, Kurt, Macho, Volker (Eds.), Forschung und wissenschaftliches Rechnen 2003. GWDG Bericht 63. Ges. für Wiss. Datenverarbeitung, Goettingen, 79-93.

Kugler, T., Kausel, E., Kocher, M., 2012. Are groups more rational than individuals? A review of interactive decision making in groups. Wiley Interdisciplinary Reviews: Cognitive Science 3 (4), 471-482. doi: 10.1002/wcs.1184.

Laughlin, P.R., 1980. Social combination processes of cooperative problem-solving groups on verbal intellective tasks, in: Fishbein, M. (Ed.), Progress in Social Psychology, Erlbaum, Hillsdale NJ, pp. 127-155.

Laughlin, P.R., 2011. Group Problem Solving. Princeton University Press, Princeton & Oxford.

Laughlin, P.R., Bonner, B.L., Miner, A.G., 2002. Groups perform better than the best individuals on Letter-to-Numbers problems. Organizational Behavior and Human Decision Processes 88, 606-620. doi: 10.1016/S0749-5978(02)00003-1.

Mesmer-Magnus, J., DeChurch, L., Jimenez-Rodriguez, M., Wildman, J., Shuffler, M., 2011. A meta-analytic investigation of virtuality and information sharing in teams. Organizational Behavior and Human Decision Processes 115, 214-225. doi: 10.1016/j.obhdp.2011.03.002.

Oechssler, J., Roider, A., Schmitz, P.W., 2009. Cognitive abilities and behavioral biases. Journal of Economic Behavior & Organization 72, 147-152. doi: 10.1016/j.jebo.2009.04.018.

Toplak, M.E., West, R.F., Stanovich, K.E., 2011. The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. Memory & Cognition, 39 (7), 1275-1289. doi: 10.3758/s13421-011-0104-1.

Weizsäcker, G., 2010. Do we follow others when we should? A simple test of rational expectations. American Economic Review 100, 2340-2360. doi: 10.1257/aer.100.5.2340.