GATE LYON SAINT-ÉTIENNE
UMR 5824

93, chemin des Mouilles
69130 Ecully - France

Maison de l'Université, Bâtiment B
10, rue Tréfilerie
42023 Saint-Etienne cedex 02 - France

http://www.gate.cnrs.fr
gate@gate.cnrs.fr

WP 2301 – February 2023

How group deliberation affects individual distributional preferences: An experimental study

João V. Ferreira, Erik Schokkaert, Benoît Tarroux

Abstract:

We study experimentally the impact of group deliberation on individual distributional preferences. We elicit subjects' distributional preferences before and after group deliberation and estimate the relative weight of persuasion, social identity, and social comparison on the effect of deliberation. We find that 10 minutes of non-binding written group deliberation has a large effect on individual (private) distributional preferences. First, post-deliberation distributional preferences are more egalitarian than pre-deliberation preferences. Second, group polarization decreases after group deliberation. Finally, we find that social identity is the main but not unique driver of this effect. Persuasion and social comparison also impact individual preferences, particularly for subjects who are not monetarily affected by the distributive outcome. Our results bring novel insights for the elicitation of distributional preferences and the design of deliberative institutions.

Keywords:
Group deliberation; Distributional preferences; Social identity; Persuasion; Social comparison

How group deliberation affects individual distributional preferences: An experimental study

João V. Ferreira[*], Erik Schokkaert[†], Benoît Tarroux[‡]

February 2, 2023

Abstract

We study experimentally the impact of group deliberation on individual distributional preferences. We elicit subjects' distributional preferences before and after group deliberation and estimate the relative weight of persuasion, social identity, and social comparison on the effect of deliberation. We find that 10 minutes of non-binding written group deliberation has a large effect on individual (private) distributional preferences. First, post-deliberation distributional preferences are more egalitarian than pre-deliberation preferences. Second, group polarization decreases after group deliberation. Finally, we find that social identity is the main but not unique driver of this effect. Persuasion and social comparison also impact individual preferences, particularly for subjects who are not monetarily affected by the distributive outcome. Our results bring novel insights for the elicitation of distributional preferences and the design of deliberative institutions.

Keywords: Group deliberation; Distributional preferences; Social identity; Persuasion; Social comparison.

[*]Corresponding author. University of Southampton, UK. E-mail: j.ferreira@soton.ac.uk; Postal address: Building 58, Highfield Campus, SO17 1BJ, Southampton, UK.

[†]Department of Economics, KU Leuven, Belgium. E-mail: erik.schokkaert@kuleuven.be; Postal address: Naamsestraat 69, B-3000 Leuven, Belgium.

[‡]University Lumière Lyon 2 and GATE Lyon-Saint-Etienne, France. E-mail: benoit.tarroux@univ-lyon2.fr; Postal address: 93, Chemin des Mouilles, 69130 Ecully, France.

1. Introduction

People's preferences about the distribution of resources can explain popular support for various tax schemes and be an important consideration for the design of redistributive policies. Popular support for redistribution is also thought to be necessary to sustain socially-oriented institutions such as the ones that make up the welfare state. An extensive literature has therefore emerged around the elicitation of distributional preferences, both in an experimental setting (e.g. Konow 2000; Cappelen et al. 2007, 2010, 2013; Balafoutas et al. 2013, 2014; Durante et al. 2014; Cettolin and Riedl 2017; Almås et al. 2010, 2020) and through large-scale survey-based experiments (e.g. Kuziemko et al. 2015; Weinzierl 2017; Alesina et al. 2018; Andreoli and Olivera 2020; Fisman et al. 2020, 2021).[1]

Although most of this literature implicitly assumes that distributional preferences are a "fixed" feature of individual preferences, most economic and political decisions with distributional consequences are preceded by a stage of group deliberation that may affect these preferences. For example, in most political systems there is a stage of public deliberation before the voting process takes place. The same holds for teams, committees, and other groups that take decisions with distributional consequences. This is particularly relevant since post-deliberation preferences are the ones that usually feed into economic and political decision-making. This means that the institutional settings that frame our social interactions may support certain preferences over others, and a better insight into the effect of group deliberation on preferences is necessary to understand the processes and outcomes of economic and political decision-making.

In this paper, we study the impact of group deliberation on individual distributional preferences. We use a lab experiment to elicit subjects' distributional (revealed) preferences before and after group deliberation and to estimate the relative weight of three of the most cited mechanisms to explain social influence in deliberative processes: persuasion, social identity, and social comparison.[2] At the beginning of the experiment, subjects are randomly assigned to a position in groups of 5 members, from "rich" to "poor", and keep this position during the

[1] See Alesina and Giuliano (2011), Schokkaert and Tarroux (2021) and Mengel and Weidenholzer (2022) for reviews.

[2] The social psychology literature has proposed three broad and dominant theories based on these mechanisms. We provide a brief overview of those theories in Section 2.

whole experiment. Their task is to choose between 7 different monetary allocations with a trade-off between efficiency (as total payoff) and equality (as, e.g., in Engelmann and Strobel 2004). This choice is individual and private (as opposed to a group choice). Increasing equality is monetarily costly for the two richest members, has no effect on the median (impartial) member, and is aligned with (monetary) self-interest for the two poorest members. In the *group deliberation treatment*, subjects are faced with this choice before (period 1) and after (period 2) 10 minutes of non-binding written group deliberation (i.e., *cheap talk*), and they are faced again with the same choice with a different group (period 3). In the *individual deliberation treatment*, after period 1 subjects are informed about the decisions that their group members took in that period, and then have 10 minutes to write a short essay on reasons for and against the different allocations. After these 10 minutes they face again the same choice in period 2 with their original group and in period 3 with a different group. As we explain below, this simple design allows us to disentangle the effects of persuasion, social identity, and social comparison.

We find that group deliberation has a statistically significant and large effect on distributional preferences at the individual and group levels. At the individual level, subjects' post-deliberation distributional preferences are statistically significantly more egalitarian than their pre-deliberation preferences. This effect is large. For example, the proportion of subjects preferring the most unequal allocation decreases from 40% before group deliberation to 20% after group deliberation, while the proportion of subjects preferring the most egalitarian allocation increases from 46% before group deliberation to 65% after it. This is the case even though the context is simple, decisions are private, the overwhelming majority of the poorest members (who benefit from equality) chose the most egalitarian allocation in all periods, and equality is monetarily costly to the two richest members. At the group level, we find that group polarization decreases after group deliberation. The decrease is large and statistically significant, irrespective of the measure we use for group polarization in preferences.

In terms of the underlying motives for these changes, our analysis suggests that social identity is the main driver behind preference change. Social identity has a statistically significant impact on distributional preferences and explains about half of the effect of group deliberation. This effect is particularly strong for the two richest members, for whom it is monetarily costly to choose more egalitarian allocations. However, social identity alone cannot explain the large effect of deliberation. The remaining half of this effect is explained by the statistically signif-

icant joint impact of persuasion and social comparison. For the median (impartial) members, social identity has no explanatory power, and persuasion seems to be the main driver behind their egalitarian shift. The relevance of persuasion in our setting is further supported by the analysis of the chat content and its relation to preference change. These results bring novel insights for the elicitation of distributional preferences and the design of deliberative institutions, which we explore below.

1.1 Relation to the literature

Our paper makes several contributions to the literature. First, our analysis is directly linked to the previously mentioned experimental literature on the elicitation of distributional preferences. To the best of our knowledge, we are the first to show in a controlled environment that *individual* distributional preferences change and are more egalitarian because of group deliberation. The most related result is from a recent paper by Ueshima et al. (2021), who show that distributional decisions for unknown others (i.e., as a spectator) are different before and after a phase of face-to-face communication *and* consensual decision-making in pairs. However, their design is not able to disentangle if the observed result on spectators' distributional decisions is due to deliberating with another person, due to being asked to reach a consensual decision after deliberation, *or* both.[3] Our results show that individual distributional preferences are not "fixed" and that group deliberation has the potential to enhance the social solidarity that is thought to be needed to sustain redistributive institutions.

Second, our decomposition of the effect of group deliberation contributes to the literature studying the underlying mechanisms of social influence in deliberative processes (e.g. Cason and Mui 1997; Frey and Meier 2004; Chen and Li 2009; Luhan et al. 2009; Chen et al. 2010; Chen and Chen 2011; Penczynski 2016). To the best of our knowledge, Cason and Mui (1997) and Luhan et al. (2009) are the only empirical studies that test the relative explanatory power of two of the mechanisms we consider (persuasion and social comparison). Cason and Mui (1997) focus on the team dictator game, where a team of two players dictates the allocation of money between them and another team of two players. They find that team choices tend to be dominated by the more altruistic member, and claim that this finding is more consistent with

[3]For a similar reason, our paper is related but significantly different from the experimental literature comparing individual and team decision-making (see Charness and Sutter 2012 and Sutter et al. 2020 for reviews). Contrary to these papers, we investigate the impact of deliberating with others on individual (private) preferences.

social comparison than with persuasion. Luhan et al. (2009) replicate their experiment using a written chat (as opposed to face-to-face communication) and find that teams are instead more selfish than individuals. In our experimental setting, we find that social identity (which they do not take into consideration) explains about half of the effect of group deliberation, with the other half explained by persuasion and social comparison. This decomposition is important for the design of deliberative institutions. For example, if social identity were the only explanation for the effect of group deliberation, then deliberating with others would not bring additional value when compared with other activities that are similarly effective at enhancing social identity (e.g. meeting with others for a fun event); but if persuasion also plays a role in shaping preferences, as our results suggest, (on-topic) group deliberation has a relevant independent effect on (distributional) preferences.

Third, our analysis contributes to the literature on polarization of opinions. In a seminal paper, Sunstein (2002, p. 178) argued that deliberation may lead to polarization as group members often "move and coalesce" toward more extreme positions that are aligned with the typical predisposition within the group. Following the large empirical and theoretical literature in social psychology on group polarization (e.g. Myers and Lamn 1976), Sunstein (2002, p. 179) pointed to persuasion and social comparison as potential mechanisms underlying this tendency. Other authors have argued in the opposite direction (e.g. List 2018), and empirical evidence from deliberative polls (e.g. List et al. 2013; Fishkin 2018) and face-to-face deliberation in the lab (e.g. Ambrus et al. 2015) support the latter hypothesis. In our experiment, we find that group deliberation leads to lower polarization in distributional settings, which had not been previously studied in this literature. Moreover, we contribute to this literature by showing that this tendency is likely driven by social identity which was not taken into account by Sunstein (2002) and others in their analysis.

Finally, our paper is related to the large experimental literature in economics on the effect of communication on preferences and behavior (e.g. Crawford 1998; Andreoni and Rao 2011; Goeree and Yariv 2011; Kittel et al. 2014; Ambrus et al. 2015; Brandts et al. 2015; Brandts et al. 2022). From these, the experiment by Ambrus et al. (2015) is the closest to ours. Their main goal is to identify the individuals who influence group decisions — in terms of the relative position of their opinion in the group —, in both gift-exchange and lottery decisions that are taken after face-to-face deliberation. In the case of gift-exchange, they are able to study the

combined effect of deliberation *and* making a group decision on individual preferences (similar to Ueshima et al. 2021). They do this within the same group and across groups (similar to what we do with periods 2 and 3). First, they show that the difference between group decisions and own decisions is correlated with group decisions within the same group, but not across groups. Consistent with our results, this suggests that social influence via deliberation is most effective within groups. Second, they observe that preferences converge toward the choices made by the members with the median individual opinion and the close-to-the-median opinion. This contrasts with our results, where preferences move away from the median individual opinion towards the most egalitarian allocation at one extreme. The experimental context may explain such a difference, since — unlike our setting — the gift exchange decision does not affect the payoff distribution within the group. Our results complement each other and contribute to the understanding of how communication influences the formation and evolution of preferences/behavior within-subjects, which could not be captured with a between-subjects design. Importantly, the between-subject element of our design allows us to reject relevant alternative explanations linked to experimental and order effects. We also contribute to this literature from a methodological point of view, by providing a novel yet simple design that can be easily replicated in the lab and in the field to study the underlying mechanisms of social influence.

The remainder of the paper is organized as follows. In Section 2, we provide a brief overview of the theories we consider to potentially underlie the effect of group deliberation. We describe our experimental design in Section 3. In Section 4, we discuss how we measure our main variables of interest and propose some hypotheses for observed behavior. Our results are summarized in Section 5. In Section 6, we discuss our results and potential alternative explanations. Section 7 concludes.

2. Conceptual background

There is an extensive literature on social influence in deliberative processes and researchers have proposed different explanations for preference change. Among these, the social psychology literature has proposed three broad and dominant theories: Persuasive Argument Theory (PAT), Social Comparison Theory (SCT), and Social Identity Theory (SIT).[4] In this section, we

[4]Other prominent theories include conformity, social pressure, and promises & threats. We return to these potential alternative explanations in Section 6.

briefly describe the main elements of these theories. Later, in Section 4, we explain how our experimental design allows us to estimate the relative weight of each of them on the effect of deliberation, and we derive some hypotheses based on the application of these theories to our setting.

According to PAT (e.g. Burnstein et al. 1973; Brown 1974, 1986), deliberation is in the first place a medium for information and argument exchange. This theory is based on the intuitive idea that individual preferences are influenced by the *most prevalent* and *convincing* arguments in a discussion.[5] PAT's explanation for preference change is therefore rooted in the exchange of arguments that happens during group deliberation. Contrary to this, the next two theories explain social influence without relation to the exchange of arguments.

According to SCT (e.g. Pruitt 1971a, 1971b; Myers and Lamn 1976), individuals want to perceive and present themselves in a socially desirable way.[6] SCT predicts that when individuals observe how others behave, they adjust their behavior in the direction of what they perceive to be the most "socially desirable" or "admirable" action (see, e.g., Myers and Lamn 1976, p. 614; Sunstein 2002, p. 179). What is the most socially desirable action is context-dependent, and below we argue that there is a clear socially desirable/admirable action in our setting. Important for our design, this theory also predicts that the "mere exposure to the preferences of others is the necessary and sufficient condition for a shift [in preferences/choices]" (Myers and Lamn 1976, p. 613; see also Cason and Mui 1997, fn. 4). Burnstein and Vinokur (1975) further observed that exposure to others' choices produced a change in preferences *only if* subjects wrote down arguments about the decision at hand. It follows that according to SCT, providing information about others' initial behavior followed by a phase of individual deliberation is on its own sufficient to change preferences.

According to SIT, group members can modify their self-identity due to changes in their emotional attachment to the group. In the original formulation of the theory (e.g. Tajfel and Turner 1979), the focus was on ingroup favoritism. However, many researchers have applied this theory more broadly, including to social dilemmas and communication (e.g. Dawes et al.

[5]There is an influential literature in economics on opinion dynamics, which usually models persuasion as strategic transmission of private information from a sender to a receiver with different interests (e.g. Crawford and Sobel 1982; Glazer and Rubinstein 2004; Kamenica and Gentzkow 2011). In our setting, however, revealing or withholding private information to change the receiver's (Bayesian) beliefs is not central.

[6]See Bonnet and Zeckhauser (2004), Frey and Meier (2004), and Chen et al. (2010) for economic experimental papers exploring social comparison and related concepts. Conformity (e.g. Bernheim 1994; Krupka and Weber 2009) is an alternative related explanation that we discuss below and test in our data.

1988; Chen and Li 2009). A prominent experimental test of this theory in economic settings is given by Chen and Li (2009). They use several minimal tasks to enhance group identity, including communication (p. 437). They find that communication via a written chat increases self-reported group attachment and has a moderate effect on behavior (pp. 450-2). Comparing treatments with and without an outgroup, they find no statistical difference in behavior or self-reported attachment, which suggests that the presence of a clear outgroup is not necessary for the social identity effect. Therefore, according to SIT social influence in group deliberation will occur via the emotional attachment to the deliberating group.

We use these insights to design an experiment in which we are able to "isolate" the effect of each of these mechanisms. At the same time, it is worth noting that our results and their implications hold irrespective of this interpretation. We return to this point in the discussion of the results.

3. Experimental design

In this section, we present the main elements of our experimental design: (i) timeline and subjects' decisions, (ii) treatments, (iii) additional measures, and (iv) experimental procedures. The instructions given to subjects for the experiment are presented in Appendix J.

3.1 Timeline and subjects' decisions

Our experimental design replicates a situation in which a group of individuals has to (re)distribute a total payoff among them. The timeline of the experiment is summarized in Figure 1. The experiment proceeds as follows. First, subjects learn about the potential payoff allocations for the group and their position in the group. They are randomly assigned to a position from "rich" to "poor" mimicking a simple "brute luck" situation in which payoffs are the result of factors beyond individual control (Dworkin 2000). They are also informed that their position is constant throughout the experiment. Next, we elicit subjects' individual distributional (revealed) preferences in three consecutive periods in which they rank seven payoff allocations for their group. Between periods 1 and 2, there is a phase of group or individual deliberation. We describe these between-subjects treatments below. Then, between periods 2 and 3, we reshuffle the groups maintaining the subjects' relative positions in the groups using a stranger matching

protocol.

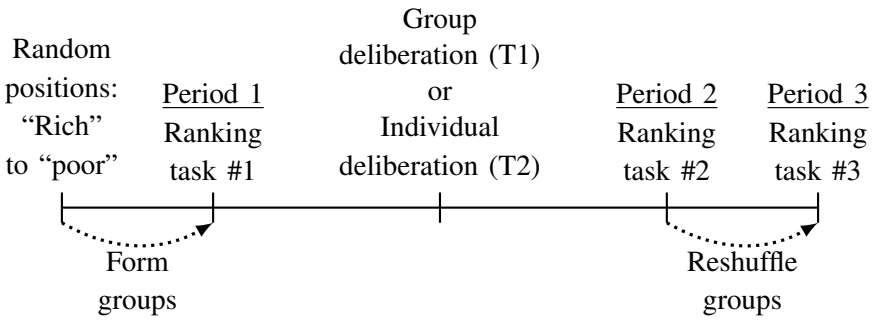

Figure 1: Timeline of the experiment

Payoff allocations. At each period, subjects rank the seven allocations of money displayed in Table 1. As it can be seen from the table, there is a trade-off between equality and efficiency (interpreted as total payoff). Allocation 1 (a1) is the most unequal and the most efficient, while allocation 7 (a7) is the most equal but the least efficient. Equality (efficiency) increases (decreases) monotonically from allocation 1 to 7. With respect to monetary self-interest, members A and B most (least) benefit from allocation 1 (allocation 7), while members D and E most (least) benefit from allocation 7 (allocation 1). Therefore, there is a conflict of interest between the "rich" (who benefit from efficiency and inequality) and the "poor" (who benefit from lower efficiency and higher equality). Member C is in an impartial position in the sense that her monetary self-interest is not affected by which allocation is chosen. Note that the differences between the allocations are not marginal. For example, a move from allocation 1 to allocation 7 involves a loss of 15.5 euros for member A and a gain of 9.5 euros for member E. To help subjects identify trade-offs, the allocations on the screen are always ordered as in Table 1 and the total payoff is presented. For similar reasons, the order of the members is always the same, from the richest (first row) to the poorest (last row). Finally, we made the subjects' own position salient by showing "me" in their row.

Ranking tasks. At each period, subjects are asked to rank the seven allocations from most preferred to least preferred and no indifferences are allowed. This choice is individual and private (as opposed to a group choice). This ranking task is completed twice in two different positions: as a dictator and as a voter.

1. **Stochastic dictator ranking task:** Subjects rank the seven allocations for the case their choice is randomly selected to determine the allocation for the group. The allocation they

	a1	a2	a3	a4	a5	a6	a7
Member A	26.5	24	21.5	19	16	14	11
Member B	16	15	14	13.5	13	12	11
Member C	11	11	11	11	11	11	11
Member D	5	6	7	8	9	10	11
Member E	1.5	3	4.5	6	8	9.5	11
Total payoff	60	59	58	57.5	57	56.5	55

Table 1: Allocations (in euros)

rank 1st has a 38% chance of being chosen, the 2nd 24%, the 3rd 17%, the 4th 11%, the 5th 7%, the 6th 3%, and the 7th 0%.

2. **Borda voting ranking task:** Subjects rank the seven allocations for the case the allocation for the group is chosen via a group vote using the Borda rule. That is, the allocation they rank 1st gets 6 points, the 2nd 5 points, ..., and the 7th none. The allocation that receives the most points is selected.

This design allows us to elicit subjects' distributional preferences in two relevant but different positions. First, the stochastic dictator mechanism gives subjects a monetary incentive to reveal an honest complete ranking of the 7 allocations. Since this ranking is kept private and subjects are effectively dictators in this position, we interpret this ranking as subjects' sincere preferences over alternatives. Second, the Borda vote allows us to elicit subjects' distributional preferences in a collective decision setting. The results may differ from the stochastic dictator choice because of different reasons, including strategic voting. In this paper, we are mainly interested in changes in distributional sincere preferences and, therefore, we focus on the stochastic dictator choices.

Incentives. At the end of the experiment, one of the three periods is randomly selected for payment. Then, for each group, one of the two decisions in that period is randomly selected for payment, with equal probabilities for the stochastic dictator and the Borda vote. In case the stochastic dictator is randomly chosen for payment, each member has a 1/5 chance of being the member who decides which allocation is chosen for payment based on her decision and the probabilities shown above. In case the Borda vote is randomly chosen for payment, the allocation that receives the most points is selected. If two or more allocations are equal at the top of the Borda ranking, the winner is the allocation (among these) that has been most often

ranked first among these allocations.

3.2 Treatments

Between periods 1 and 2, there is a phase of group deliberation (Treatment 1 or T1) or individual deliberation (Treatment 2 or T2). In the initial instructions, we announced to subjects that they would be making decisions in three independent periods, but we only informed subjects about the group/individual deliberation phase after period 1 was finished. This information flow was implemented to avoid subjects forming expectations that could affect their choices in period 1.

- **T1 (Group deliberation treatment):** Subjects have 10 minutes of non-binding text communication to discuss the allocations.

A few features of T1 are worth explaining. First, the communication was public (i.e., there was no private communication between subgroups of members as, e.g., in Goeree and Yariv 2011 or Brandts et al. 2022) and the members were identifiable by their position in the group (A to E, as opposed to being fully anonymous as, e.g., in Ahn et al. 2018 and Heap et al. 2020). Second, subjects were encouraged to discuss about the available allocations, but there were no restrictions on the content of the communication. The exact wording that appeared on the screen was: "*Before making your choices, you have the opportunity to discuss the available allocations*".[7] Third, the discussion was done via a written chat (as opposed to face-to-face as, e.g, in Ambrus et al. 2015 and Ueshima et al. 2021) and the decisions were made privately (as opposed to jointly as in Ueshima et al. 2021 or individually but publicly known as, e.g., in Cason and Mui 1997). Therefore, communication is not binding in our setting and can be seen as a form of cheap talk. Overall, our experimental group deliberation has many features of online deliberation platforms (e.g. forums) that are increasingly popular but is further removed from usual forms of face-to-face public deliberation. This design choice is important for greater experimental control as it preserves personal anonymity and avoids confounding factors such as prior acquaintance of group members unknown to the experimenter. Other design choices, such as not allowing for private communication and allowing for identification of members in terms

[7]Subjects were also asked to sign in a paper that they would respect the following minimal charter: "*The content of the discussion is not restricted. However, you are not allowed to reveal personally identifiable details such as name, surname, field of study, computer number to which you are assigned, etc*."

of their position in the group, are intended to make our experimental group deliberation closer to an "ideal" setting for public discussion without compromising high experimental control.[8]

- **T2 (Individual deliberation treatment):** Subjects are informed about the decisions that their group members took in period 1 and have 10 minutes to write a short essay on reasons for and against the different allocations.

According to SCT, including information about the initial decisions of others followed by individual deliberation is necessary and sufficient for social comparison to be at play. T2 is designed to "isolate" the effect of social comparison since persuasion and social identity cannot be at play in this treatment.

Two features of T2 are worth explaining. First, the exact wording that appeared on the screen was: "*The purpose of this essay is to help you to think about the different possible points of view. To this end, we ask you to write a text presenting arguments for or against each possible payoff distribution. These arguments could be presented by yourself or any other member of the group.*" This framing was provided following the hypothesis that social comparison is more effective when people are stimulated to think about arguments that others might have had for their choices (see Myers and Lamn 1976, p. 615). It is also intended to give a clear objective to the essay such that it feels more natural to the subjects in our experiment. We do not exclude, however, the possibility that this framing may impact behavior. Therefore, the effect we attribute to social comparison includes the effect of the information about others, individual deliberation, *and* the effect of this framing.

Second, the information given to subjects was the number of group members who ranked each allocation as their preferred option in the stochastic dictator task of period 1. This information remained visible on the subjects' screens while they wrote their essay. The reason underlying this design choice is two-fold. First, it allows for a certain symmetry between the two treatments in terms of information. Indeed, in T1 we expected subjects to infer others' preferred options via the chat but not necessarily their full ranking of allocations. Second, we presented group members' preferred allocations in one position (as opposed to, for example,

[8]We use "ideal" in the sense of a communicative procedure that "is designed to promote substantive, balanced, and civil discussion" (List 2018). In Appendix H, we report the views of subjects about the quality of deliberation that we recovered in a post-experimental questionnaire and show that this has an impact on the effect of group deliberation on distributional preferences.

the average/median choices as in Bonnet and Zeckhauser 2004, Frey and Meier 2004, or Chen et al. 2010) such that subjects could infer, without being overwhelmed with information, the choices made by different members in their group. Krupka and Weber (2009) use a similar informational procedure in the dictator game and find that observing others being generous significantly increases pro-social behavior, especially when the group's norm (in terms of numerical majority) is apparent. In Appendix F, we test and reject the hypothesis that conformity to the group's norm explains our results.

3.3 Additional measures

We recorded the chat communication in T1, which we analyze in Appendix E to provide further support to some of our hypotheses and results. Before the lab session, we also gathered information on personality traits via a short online survey. This survey consisted of three influential psychometric questionnaires: Big Five (BFI-S), (social) Open mindedness and Machiavelism. Finally, at the end of T1's lab session, we asked several questions about the perceived quality of the deliberation. Questions included subjects' perception of equality of speech, respect, the willingness of members to justify their views, and others' sincerity. We explore the relationship between personality traits, the perceived quality of deliberation, and the effect of group deliberation in Appendix H.

3.4 Procedures

The experimental sessions took place in the experimental lab of the University of Lyon (GATE-LAB), France, in 2019. We conducted 10 sessions with a total of 240 participants, 175 in T1 and 65 in T2. We conducted 7 sessions for T1 and 3 sessions for T2 with 20, 25, or 30 participants. The sessions lasted around 1 hour and the mean payment was 22 euros with a 7 euros show-up fee.

A large majority of our subject pool are students (89% in T1, 94% in T2) from economics and management (66% in T1 and 72% in T2) and engineering (22% in T1 and 17% in T2). Women represent 48% of subjects (49% in T1 and 48% in T2) and the mean age was 23 in both treatments. Perceived social categories (from rich to poor) are also similar across the two treatments, and we find no statistically significant difference across treatments on this and other social-demographic observed characteristics. Table A14 in Appendix G summarizes this

information.

4. Measurement and hypotheses

4.1 Individual degree of egalitarianism

For our analysis, we focus on what we call the subjects' *degree of egalitarianism*. Let us represent the degree of egalitarianism of subject i at period $t \in \{1,2,3\}$ in treatment $d \in \{1,2\}$ by $\alpha^t_{i,Td}$. We shorten this notation to α^t_i whenever the treatment is irrelevant and to α^t_{Td} when referring to average treatment effects. In our experimental setting, the allocations are set up such that if allocation k has higher equality than allocation k' then allocation k has lower total payoff than allocation k'. This represents the potential trade-off between equality and total output. It follows that one can interpret $\alpha^2_i > \alpha^1_i$ as an indication that i in period 2 is more concerned with equality *and/or* less concerned with total payoff than she was in period 1. For the two richest members, monetary self-interest and equality are *not* aligned. Therefore, a higher degree of egalitarianism has an additional cost for these members, and $\alpha^2_i > \alpha^1_i$ for a rich member would also suggest that this subject is less concerned with monetary self-interest in period 2 than in period 1. For the median (impartial) member, monetary self-interest is not at play. Changes in preferences are driven by other-regarding concerns over equality and total payoff. For the two poorest members, a higher degree of egalitarianism can be driven by monetary self-interest. This means that a greater concern for equality may not be the underlying reason for $\alpha^2_i > \alpha^1_i$. However, as shown below, the overwhelming majority of "poor" members choose — not surprisingly — the allocation that benefits them most/has greater equality already in period 1, and changes in α^t_{Td} are driven by rich and median members. "Degree of egalitarianism" is therefore a useful simplification, but our interpretation of the results takes the interplay of equality, efficiency, and self-interest into play.[9]

There are several possible proxies for the degree of egalitarianism of individual subjects. In our main analysis, we use two proxies. The simplest one is the *preferred allocation*. A subject is considered more egalitarian than another if her preferred allocation is more egalitarian than the preferred allocation of the other. Similarly, a subject is more egalitarian in one period than

[9]Maximizing equality and the minimum payoff are also perfectly aligned in our setting. This means that we cannot distinguish between these two motivations and a higher degree of egalitarianism can be driven by a greater concern for the least well-off.

in another if her preferred allocation is more egalitarian in the former than in the latter. While appealing for its parsimony, this first proxy ignores the overall ranking given by subjects.

A more general picture is offered by our second proxy, which is the *Kemeny distance* (or K-distance) between a subject's rankings and the "least egalitarian ranking" (i.e., the ranking in which the least egalitarian allocation is the preferred allocation, the second least egalitarian allocation is the second preferred, and so on).[10] Formally, we consider the 7×7 matrix of subject's i pairwise comparisons of allocations k and k', $Q_i = (q_i^{kk'})$ where:

$$q_i^{kk'} = \begin{cases} 1 & \text{if } i \text{ prefers } k \text{ to } k' \\ 0 & \text{if } i \text{ is indifferent between } k \text{ and } k' \\ -1 & \text{if } i \text{ prefers } k' \text{ to } k \end{cases}$$

The matrix $R = (r^{kk'})$ corresponding to the least egalitarian ranking is similarly defined. The Kemeny distance can then be written as follows:

$$d(Q_i, R) = \sum_k \sum_{k'} |q_i^{kk'} - r^{kk'}| \tag{1}$$

The larger the value of $d(Q_i, R)$, the higher the degree of egalitarianism of subject i. We normalize the distance measure such that the degree of egalitarianism lies between 0 (no distance between the subject's ranking and the least egalitarian ranking) and 1 (maximal distance between the subject's ranking and the least egalitarian ranking).

In Appendix C, we show that the degree of egalitarianism α_i^t can also be interpreted as a parameter of a utility function with social preferences. The results using this approach are similar to the other proxies.

4.2 Persuasion, social comparison, and social identity

The total effect of group deliberation on the degree of egalitarianism, as derived from the results in treatment 1, can be written as follows:

$$\alpha_{i,T1}^2 - \alpha_{i,T1}^1 \tag{2}$$

[10] The Kemeny (1959) distance, also known as Kendall-tau distance or swap distance, among other names, is one of the most used distance functions for comparing ordinal preferences. However, our analysis can also be done with other distance functions between ordinal preferences. See Can (2014) for a review.

Our experimental design with two treatments and three periods allows us to measure the relative explanatory power of persuasion, social identity, and social comparison on this total effect. To see this, Table 2 indicates if each of these mechanisms can influence decisions in a particular period and treatment. In period 1 (equal in both treatments), individuals make their decisions in private and without any interaction with others or information on their decisions. For this reason, none of the mechanisms can be at work. In the second period of T1, the three mechanisms are potentially at work, summing to the total effect of group deliberation captured in expression (2). In the third period of T1, subjects are allocated to a group with whom they have not interacted. For this reason, the social identity effect should not be at work. This means that if social identity was the only explanation for social influence, then preferences should change between periods 1 and 2 but they should return to pre-deliberation levels in period 3 as there is no emotional attachment to that group. On the contrary, if persuasion and social comparison were the only mechanisms at work, period 2 and period 3 preferences should be the same as these effects should not disappear when groups are reshuffled. In the second period of T2, subjects have been informed about the initial decisions of others and wrote an essay about reasons for and against the different allocations. According to SCT, this is necessary and sufficient to induce social comparison during individual deliberation. Besides that, in T2 individuals do not interact, so the channels of social identity and persuasion are turned off.

		Treatment 1 Group deliberation		**Treatment 2 Individual deliberation**	
	Period 1	**Period 2**	**Period 3**	**Period 2**	**Period 3**
Social identity	no	yes	no	no	no
Social comparison	no	yes	yes	yes	yes
Persuasion	no	yes	yes	no	no

Notes: The "yes" and "no" in the cells indicate if a given mechanism can influence decisions in a particular period and treatment.

Table 2: Decomposition of the effect of group deliberation

If we make the (untestable) assumption that the effects of the three mechanisms are additive, this simple design allows us to estimate the weight of each of them in the effect of group

deliberation on distributional preferences:

$$\begin{aligned} \textbf{Social identity} &= \alpha^2_{i,T1} - \alpha^3_{i,T1}. \\ \textbf{Social comparison} &= \alpha^2_{i,T2} - \alpha^1_{i,T2} \equiv \alpha^3_{i,T2} - \alpha^1_{i,T2}. \\ \textbf{Persuasion} &= (\alpha^3_{i,T1} - \alpha^1_{i,T1}) - (\alpha^3_{i,T2} - \alpha^1_{i,T2}) \equiv \alpha^3_{i,T1} - \alpha^3_{i,T2}. \end{aligned} \tag{3}$$

The sum of these effects is equal to the total effect of group deliberation in expression (2) above.

4.3 Group polarization

We measure group polarization in preferences in two ways. First, we use the polarization measure proposed by Esteban and Ray (1994). Assume that the degree of egalitarianism can take q values and write these values as $(\overline{\alpha}_1, ..., \overline{\alpha}_q)$. In our setting, the vector $\overline{\alpha}$ corresponds to a number of discrete values for the Kemeny distance measure. The Esteban-Ray measure of polarization at period t in treatment d can then be written as follows:

$$P^{\gamma,t}_{Td} = \sum_{l=1}^{q} \sum_{m=1}^{q} (\pi^t_{l,Td})^{1+\gamma} \pi^t_{m,Td} |\overline{\alpha}_l - \overline{\alpha}_m| \tag{4}$$

where $\pi^t_{l,Td}$ is the proportion of subjects with a degree of egalitarianism $\overline{\alpha}_l$ at period t in treatment d. The parameter γ can be chosen and indicates the weight given to polarization as compared to inequality. In fact, if $\gamma = 0$, expression (4) reduces to the traditional Gini-coefficient (up to a normalization factor). To see the intuition behind this measure of polarization, assume that there are two "subgroups" within a deliberating group in our setting, one "egalitarian" and one "non-egalitarian". According to this measure, the closer the preferences between these two subgroups are, the lower the polarization. At the same time, this measure also implies that the more homogeneous the two subgroups are, the higher the polarization.

Second, we use the polarization notion proposed by Sunstein (2002, p. 178), according to which polarization increases as groups converge towards more extreme positions that are aligned with the typical predisposition within the group. According to this notion, group polarization increases if groups with an initial median preferred "non-egalitarian" allocation (e.g. allocation 1 or 2) converge towards allocation 1, and groups with an initial median preferred

"egalitarian" allocation (e.g. allocation 6 or 7) converge towards allocation 7. Note that in such circumstances, polarization, as measured by Esteban and Ray (1994), would decrease. In Appendix D, we show that our results also hold for the Herfindahl index of polarization.

4.4 Hypotheses

We can make a few hypotheses about the expected behavior in our experiment by applying PAT, SCT, and SIT to our setting. Take SCT first. We hypothesized that subjects perceived the egalitarian allocation as the most socially desirable action. While we cannot test this auxiliary hypothesis in our setting, evidence from standard dictator games (and dictator games in which dictators "take" instead of "give" money to the recipient) shows that "[t]here is substantial social agreement that the action that produces equal payoffs is very socially appropriate" (Krupka and Weber 2013, p. 506). Conversely, dictators maximizing their payoff is considered to be the most socially inappropriate action (Krupka and Weber 2013, p. 506). This evidence suggests that egalitarian allocations are socially desirable/accepted in situations where relative positions are the result of brute luck as in our experiment. Therefore, we predicted social comparison to shift preferences in the egalitarian direction.

Consider now PAT. Since the two poorest members benefit from equality, we anticipated at least two initial positions to be egalitarian. This is relevant since PAT predicts that "discussion generates arguments predominantly favoring the initially preferred alternative, including some persuasive arguments that the typical subject has not previously considered" (Myers and Lamn 1976, p. 611). In addition, given that in our setting subjects are randomly assigned to their positions, we also expected arguments in favor of the egalitarian allocation to dominate the discussion and to be easier to justify than arguments in favor of non-egalitarian allocations. The evidence just cited on the social desirability of equality in similar settings provides support to this auxiliary hypothesis, which we (partially) test in Appendix E using chat content analysis. Therefore, we predicted that egalitarian arguments would be more prevalent and persuasive, which would further increase the observed average degree of egalitarianism after deliberation.

Finally, consider SIT. In our setting, communicating with others within a group is likely to promote emotional attachment to that group. According to SIT, this is likely to lead individuals to minimize in-group inequalities. Costa-Font and Cowell (2015) survey the literature on the link between social identity and distributional preferences and conclude that the evidence

suggests that group identity gives rise to welfare-maximizing actions and pro-distributional social preferences. We therefore predicted that social identification would also change subjects' preferences towards more egalitarian allocations. It follows that we can make the following hypotheses:

Hypothesis 1 Social comparison increases the average degree of egalitarianism ($\alpha^3_{T2} - \alpha^1_{T2} > 0$).

Hypothesis 2 Persuasion increases the average degree of egalitarianism ($\alpha^3_{T1} - \alpha^3_{T2} > 0$).

Hypothesis 3 Social identity increases the average degree of egalitarianism ($\alpha^2_{T1} - \alpha^3_{T1} > 0$).

If at least one of these mechanisms is at play in group deliberation, and no mechanism works in the opposite direction as predicted, we can derive immediately:

Hypothesis 4 The average degree of egalitarianism increases after group deliberation ($\alpha^2_{T1} - \alpha^1_{T1} > 0$).

Since only SCT is relevant in the case of individual deliberation, we can also put forward:

Hypothesis 5 The average degree of egalitarianism increases after individual deliberation, but less so than after group deliberation ($\alpha^2_{T1} - \alpha^1_{T1} > \alpha^2_{T2} - \alpha^1_{T2} > 0$).

Finally, there is conflicting evidence on the effect of deliberation on group polarization and little evidence in settings similar to ours. However, we conjectured that polarization would decrease after group deliberation in our setting. We expected poor subjects to select the most egalitarian allocation before and after group deliberation (which is in their self-interest), while rich and median subjects would shift towards more egalitarian allocations after group deliberation due to persuasion, social comparison, and social identity. These patterns of behavior would lead to lower polarization:

Hypothesis 6 The polarization of subjects' degree of egalitarianism decreases after group deliberation ($P^{\gamma,2}_{T1} - P^{\gamma,1}_{T1} < 0$).

5. Results

Unless otherwise stated, all the results in this section refer to the stochastic dictator task. The results with the Borda vote are presented in Appendix I.

	Mean	% with K-distance			Signed-rank tests	
		$=0$	$\in]0,1[$	$=1$	= period 2	= period 3
Period 1	0.52	33.71	23.43	42.86	$p<0.001$	$p=0.034$
Period 2	0.71	17.14	28.57	54.29		$p=0.048$
Period 3	0.62	29.14	18.86	52.00		

Notes: Signed-rank test for clustered data, clustering at the group level for p1=p2 and at the session level for p2=p3 and p1=p3 (see Datta and Satten 2008).

Table 3: Kemeny distances across periods in T1

5.1 The effect of group deliberation on individual distributional preferences

The distribution of preferred allocations over the three periods in our group deliberation treatment is given in Figure 2. In all periods, the overwhelming majority of subjects preferred either the most unequal allocation (a1) or the most egalitarian allocation (a7). This is not surprising, as these allocations align with the monetary self-interest of rich and poor members respectively. However, the proportion of subjects preferring allocation 1 decreases from 40% before group deliberation (first period) to 20% after group deliberation (second period). At the same time, the most egalitarian allocation is more attractive after deliberating with others: while 46% of subjects prefer allocation 7 in the first period, this proportion increases to 65% in the second period. In the third period, i.e., after reshuffling the groups, we see an "intermediate" distribution with 32% of subjects selecting allocation 1 as their preferred one and 55% selecting allocation 7. Non-parametric tests indicate that the shifts in the preferred allocations between periods 1 and 2 (from a1 to a7) and the reverse shift between periods 2 and 3 are statistically significant ($p<0.01$ for allocations 1, 2, and 7, Stuart-Maxwell tests of homogeneity; see Table A3 in Appendix A for the full set of tests). Although less outspoken, they also suggest that there remains a statistically significant difference between the preferred distributions in periods 1 and 3 ($p<0.05$ for allocations 1, 2, and 7, Stuart-Maxwell tests of homogeneity).

A similar pattern is observed in Table 3 for the Kemeny distance measure. Preferences become more egalitarian after group deliberation. This change is large and statistically significant. After reshuffling the groups, there is some return to the original less egalitarian positions. However, this return is not complete and there remains a statistically significant difference between periods 1 and 3.

The average results presented so far are further supported when looking at changes at the individual level. Let us focus on the Kemeny distance (the results with the preferred allocation

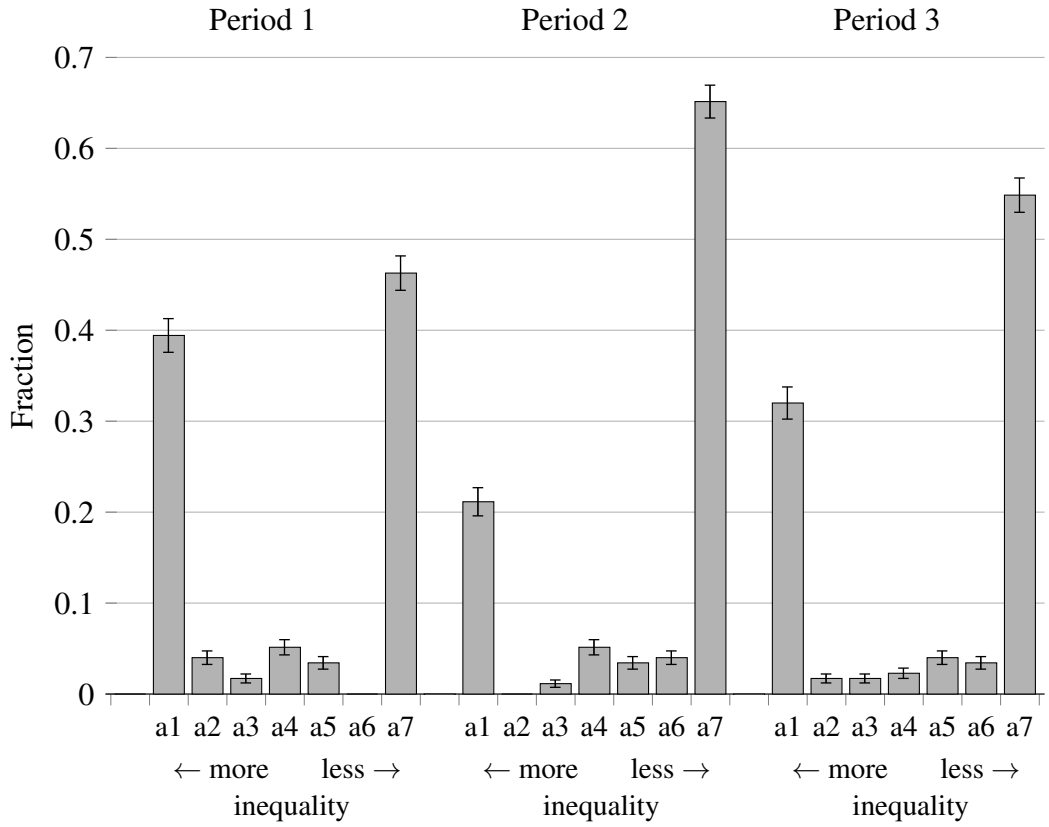

Figure 2: Distribution of preferred allocations in T1 across periods

Notes: The main bars display percentages of subjects with preferred allocations a1 to a7 per treatment. The error bars display one standard error of the mean, with values based on logit estimates with dummies for allocations for each period.

are equivalent). Table 4 shows that individual preferences do change after group deliberation and that a majority of subjects are more likely to support more egalitarian allocations after deliberation. There is clear evidence that the least egalitarian members in the first period tend to become more egalitarian in period 2. We find that 32 of the 59 subjects (54%) preferring the least egalitarian ranking in the first period (K-distance $= 0$, see the top left row of the table) opt for a more egalitarian ranking in the second period, and 9 of these (15%) shift for the most egalitarian ranking (K-distance $= 1$). In period 3, some subjects return to their initial preference. For example, while 32 subjects preferring the least egalitarian ranking in the first period opt for a more egalitarian ranking in the second period, only 13 subjects select a more egalitarian allocation in period 3 than in period 1 (see the top right row of the table).

The degree of preference shifting depends on the randomly assigned position of individuals. Results are shown in Table 5. For the two members with the lowest incomes across allocations, considerations of self-interest and equality coincide. Not surprisingly, the overwhelming majority of these subjects (78.57%) opt for the most egalitarian ranking in period 1, and we observe

Period 1	Period 2				Period 3			
	$=0$	$\in]0,1[$	$=1$	All	$=0$	$\in]0,1[$	$=1$	All
$=0$	27	23	9	59	46	10	3	59
$\in]0,1[$	3	19	19	41	5	20	16	41
$=1$	0	8	67	75	0	3	72	75
All	30	50	95	175	51	33	91	175
Test SM and MB:	$p < 0.001$ and $p < 0.001$				$p = 0.001$ and $p = 0.004$			

Notes: This table reports the number of subjects with K-distance equal to 0 (least egalitarian ranking), 1 (most egalitarian ranking), and in between across periods. SM stands for the Stuart-Maxwell test of marginal homogeneity and MB for the McNemar-Bowker test for symmetry.

Table 4: Individual changes in Kemeny distances

a small egalitarian shift over time that is only statistically significant between periods 1 and 3. The median (impartial) member faces a trade-off between other-regarding concerns over equality and total payoff. As Table 5 shows, equality considerations dominate already in period 1 before group deliberation, with a mean Kemeny distance of 0.69 to the least egalitarian ranking. Still, this egalitarian bent gets considerably stronger after deliberation ($p = 0.022$, signed-rank test with clustering at the group level). Importantly, this effect is not reversed in period 3. In other words, contrary to the average picture, for impartial members the egalitarian shift due to group deliberation carries over to other groups. Things look very different for the richest group members, who face a conflict between self-interest/efficiency and egalitarianism. In this setting, they are very self-interested/efficiency-minded before group deliberation, with 75.71% choosing the most inegalitarian ranking (most efficient/in their monetary self-interest). Yet, also for them, deliberation has a large and statistically significant egalitarian effect ($p < 0.001$, signed-rank test with clustering at the group level). However, contrary to the other members, there is a return to inegalitarian rankings in period 3 after reshuffling the groups. This shift is statistically significant and reverses most of the effect of group deliberation for these members.

5.2 Decomposition of the effect of group deliberation

The decomposition results are shown in Table 6 (see Appendix A for the underlying regressions). First, we find that around half of the total effect of group deliberation is explained by social identity (53% with the preferred allocation, 47% with the Kemeny distance). The effect of social identity is highly significant ($p < 0.001$, Wald test). Second, we find that even though the effects of persuasion and social comparison are not statistically significant individually, the

	Mean	% with K-distance			Signed-rank tests	
		$=0$	$\in]0,1[$	$=1$	= period 2	= period 3
Rich subjects						
Period 1	0.08	75.71	22.86	1.43	$p<0.001$	$p=0.269$
Period 2	0.39	40.00	41.43	18.57		$p=0.017$
Period 3	0.16	70.00	22.86	7.14		
Median subjects						
Period 1	0.69	11.43	34.29	54.29	$p=0.022$	$p=0.051$
Period 2	0.87	2.86	25.71	71.43		$p=0.593$
Period 3	0.85	2.86	28.57	68.57		
Poor subjects						
Period 1	0.89	2.86	18.57	78.57	$p=0.208$	$p=0.032$
Period 2	0.95	1.43	17.14	81.43		$p=0.110$
Period 3	0.96	1.43	10.00	88.57		

Notes: Signed-rank test for clustered data, clustering at the group level for p1=p2 and at the session level for p2=p3 and p1=p3.

Table 5: Kemeny distances across periods per rank groups

	Total	Identity	Social Comp.	Persuasion	SC+P
Pref. allocation (coef)	0.904	0.483	0.091	0.330	0.421
Wald tests (p-values)	0.000	0.000	0.697	0.270	0.000
K-distance (coef)	0.193	0.091	0.037	0.065	0.102
Wald tests (p-values)	0.000	0.000	0.462	0.337	0.000

Notes: This table reports the decomposition results using the regression estimates reported in Table A1 (columns 2 and 4) in Appendix A. SC+P stands for the joint effect of social comparison and persuasion.

Table 6: Decomposition results

sum of the two components is found to be highly significant ($p<0.001$, Wald test). This sum corresponds to what is left when we subtract social identity. The persuasion effect is stronger than the social comparison effect but, given that these effects are estimated with a large margin of error, this conclusion can only be drawn cautiously. Our results with α_i^t being a parameter of a utility function with social preferences, reported in Appendix C, support these inferences.

Additional insights are obtained when looking at the results for subgroups of subjects (see Table 7). When looking at the two richest members for which equality is (monetarily) costly, we find that social identity explains around three-quarters of the effect of group deliberation (77% with the preferred allocation, 73% with the Kemeny distance). We also find a statistically significant joint effect of persuasion and social comparison on the Kemeny distance, but it is very difficult to differentiate them statistically. It seems, however, that persuasion is more important than social comparison for rich members. These inferences are once again supported

	Total	Identity	Social Comp.	Persuasion	SC+P
Rich subjects					
Pref. allocation (coef)	2.000	1.540	-0.379	0.839	0.460
Wald tests (p-values)	0.000	0.000	0.550	0.245	0.162
K-distance (coef)	0.311	0.228	0.015	0.068	0.083
Wald tests (p-values)	0.000	0.000	0.758	0.302	0.018
Median subjects					
Pref. allocation (coef)	1.594	0.429	0.741	0.424	1.165
Wald tests (p-values)	0.000	0.116	0.195	0.502	0.003
K-distance (coef)	0.217	0.020	0.061	0.135	0.197
Wald tests (p-values)	0.000	0.510	0.428	0.127	0.000
Poor subjects					
Pref. allocation (coef)	0.287	-0.199	1.168	-0.682	0.486
Wald tests (p-values)	0.501	0.594	0.112	0.418	0.227
K-distance (coef)	0.062	-0.011	0.047	0.026	0.073
Wald tests (p-values)	0.028	0.518	0.027	0.301	0.003

Notes: This table reports the decomposition results using the regression estimates reported in Table A2 for rich, median, and poor subjects (Appendix A). SC+P stands for the joint effect of social comparison and persuasion.

Table 7: Decomposition results per rank groups

by our results with α_i^t being a parameter of a utility function with social preferences.

The pattern is different for median members. In their case, the effect of social identity is not statistically significant. Instead, the explanation behind the statistically significant effect of group deliberation is the joint effect of persuasion and social comparison. Finally, when looking at the two poorest members, the effect of social identity is again not statistically significant. The joint effect of persuasion and social comparison is statistically significant for the Kemeny distance, and for these members social comparison seems to explain most of their small egalitarian shift.

5.3 Group polarization

We now consider the effects of deliberation at the group level. In particular, we study whether there is a move toward consensus or towards polarization. Table 8 summarizes our findings for the Esteban-Ray measure of polarization. We find that in period 2 there is a statistically significant decrease in the degree of polarization for different values of γ, including the boundary case of the Gini-coefficient ($p < 0.001$ for all cases, signed-rank tests with data clustered by group). This accords with our previous findings, according to which group members (independent of their positions) show an egalitarian shift after deliberation.

	Period 1	Period 2	Period 3
Gini $\gamma = 0$	0.45	0.26	0.37
Polarization $\gamma = 0.5$	11.31	7.83	10.78
Polarization $\gamma = 1$	13.46	9.39	13.22

Notes: This table reports the polarization of individual Kemeny distances within groups according to the measure proposed by Esteban and Ray (1994).

Table 8: Polarization within groups

We also do not find a phenomenon of polarization as described by Sunstein (2002). In fact, no matter what is the median preferred allocation of the group in period 1, all groups tend on average to become more egalitarian after group deliberation. Contrary to Sunstein's (2002) hypothesis, groups that in period 1 have a median preferred allocation of 1, 2 or 3 (the least egalitarian allocations) still increase their mean K-distance to the least egalitarian ranking in period 2. In sum, we observe a consistent shift towards the most egalitarian ranking, as opposed to a phenomenon where individuals converge to more extreme allocations that are aligned with the initial typical predisposition within the group. The results for the Herfindahl index, shown in Appendix D, are also consistent with the decrease of polarization after group deliberation.

The results for period 3 are also interesting. Take the Esteban-Ray measure of polarization. Even though there remains a statistically significant decrease in group polarization between periods 1 and 3 ($p < 0.001$ for all cases, signed-rank tests with data clustered by session), the average level of polarization returns to values close to the ones observed before deliberation. If we interpret this result in line with the approach that was applied in the previous subsection, it suggests that the building up of social identity is the dominant social mechanism behind the decrease in polarization. This may help interpret some of the conflicting results that have been found in previous studies, since predictions of polarization through group deliberation usually emphasize persuasion and social comparison as the main driving forces behind that polarization.

6. Discussion

We start our discussion by interpreting our results without reference to social influence mechanisms. First, we observe a clear egalitarian shift away from efficiency and lower polarization after group deliberation on the distribution of resources within a group. We are the first to show

this in a controlled environment. Second, the comparison of the preferences in T1 before and after reshuffling the groups (periods 2 and 3) can tell us how much of the impact of group deliberation carries on to other groups. In our setting, the effect carries over for impartial members but not so much for members who went against their personal interests after group deliberation. Third, the comparison between the two treatments tells us what group deliberation brings beyond individual deliberation knowing the initial preferences of others. Our results for T2, reported in Appendix B, show that subjects become more egalitarian after individual deliberation. However, this effect is much smaller in magnitude than in T1 (0.05 shift in K-distance compared to 0.19 in T1; difference is statistically significant, $p = 0.007$ and $p = 0.020$ for Wald tests based on columns 2 and 4 of Table A3 respectively) and is only statistically significant for the Kemeny distance of impartial members ($p = 0.045$, signed-rank test with clustering at group level). This suggests that individual deliberation has a place in shaping preferences, but group deliberation can have a larger independent effect.

In this paper, we have used three social influence mechanisms to interpret these observed "facts". We hypothesize that group deliberation involves these three mechanisms. The analysis of the content of the chat scripts brings additional insights into the psychological mechanisms at work (analysis reported in Appendix E). Three research assistants read all the chats and were asked to classify and code the messages according to 24 (non-exclusive) categories. We assigned a message to a specific category if at least 2 out of the 3 coders agreed. In terms of content, the chat scripts show that all groups engaged in discussion about the allocations and the deliberation was active with an exchange of (reasoned) arguments in favor or against different allocations, simple responses and propositions, greetings, goodbyes, and other out-of-topic messages. Consistent with the PAT hypothesis, the number of (reasoned) arguments has a statistically significant effect on subjects' degree of egalitarianism. Also consistent with this mechanism, pro-social arguments about equality and fairness have a positive effect on the propensity of subjects to adopt egalitarian preferences, while the effect of efficiency arguments goes in the opposite direction. The only other messages to have a consistent statistically significant effect are arguments seeking a compromise, which have a negative effect on subjects' degree of egalitarianism, suggesting that compromise is sought by rich members. These results support the insight that group deliberation affects preferences beyond social identity, and that persuasion plays an important role in that effect.

The psychological and economic literatures have proposed alternative mechanisms to interpret the effect of group deliberation. One of the most prominent is the desire to conform to the behavior that is most common in a given group (e.g. Akerlof 1980; Bernheim 1994; Krupka and Weber 2009; Chen et al. 2010). According to this theory, if given information about others' behavior, people tend to copy the behavior that is prevalent in a group (e.g. Akerlof 1980; Bernheim 1994). If this holds in our setting, we should observe a similar phenomenon as in Krupka and Weber (2009), where pro-social choices increased as subjects observed more pro-social behavior on the part of others (p. 313). Under this interpretation, conformity is an alternative explanation to social comparison. Fortunately, it is possible to distinguish both empirically. According to SCT, shifts go in the direction of the socially desirable allocation (which we have argued is the most egalitarian allocation in our setting), while according to conformity shifts go in the direction of the most prevalent behavior within the group as inferred by the numerical majority or median of the group. To distinguish between the two, we can therefore check whether groups converge to a non-egalitarian position if the numerical majority/median of the group is not egalitarian at the start of the deliberation. Contrary to conformity's prediction, we observe a systematic convergence towards the most egalitarian allocation no matter the group's median preferred allocation in the first period, in both T1 and T2. We also do not find evidence that egalitarian choices increase as subjects observed more egalitarian choices on the part of others. We report these results in Appendix F.

A related alternative explanation is social pressure exerted by group members for others to behave in a certain way. From a theoretical perspective, this does not seem to be a very powerful mechanism in our setting, since decisions are private and not disclosed to others, there is no opportunity for positive or negative reciprocity after the group deliberation, and the deliberation is through a written and anonymous chat without face-to-face contact. For the same reason, "credible" promises and threats during the deliberation are unlikely to be an important force (cf. Ellingsen and Johannesson 2004). Our chat analysis, which shows that compromise messages (including promises/threats) are rare in groups, supports these claims. Still, we do not exclude the possibility that discussing with others may generate a kind of "psychological contract" between participants (see Frey and Bohnet 1997). Under our framework, we would interpret such a psychological contract as a form of social identity.

Finally, it would be possible that elements of our design could bias our results in certain

directions. For instance, there is the possibility of experimenter demand effects (Zizzo 2010). While previous papers suggest that these effects may bias choices in period 1 towards egalitarianism (see, e.g., List 2007), they could not explain the shift between periods 1 and 2. This is the case since the framings of individual and group deliberations are not biased in a given (more or less egalitarian) direction. Another possibility is (within-subject) order effects. According to this explanation, the observed difference between periods is an artifact from the within-subjects design with several consecutive periods. While we do not exclude the possibility that decisions in periods 2 and 3 may be influenced by decisions taken in previous periods, order effects could not explain the large and statistically significant differences that we observe between T1 and T2. This means that they cannot be a significant underlying force behind the shifts we observe across periods in T1.

7. Concluding remarks

We use a lab experiment to elicit individuals' distributional preferences before and after group deliberation. Subjects' post-deliberation distributional preferences are significantly more egalitarian than their pre-deliberation preferences. Polarization decreases. We find that social identity explains about half of the effect of group deliberation, while the remaining half is explained by the joint impact of persuasion and social comparison. The effect of social identity is particularly strong for the members for whom it is monetarily costly to choose more egalitarian allocations, while the effects of persuasion and social comparison are mainly relevant for the members for whom all allocations provide the same monetary payoff. We find no evidence for alternative mechanisms such as conformity or order effects.

These results bring novel insights into distributional preferences and the design of deliberative institutions. First, our results suggest that allowing people to engage with each other and exchange arguments about the distribution of resources can help sustain pro-social behavior and institutions. This might explain why political decision-making about redistribution in the real world is more stable than one would expect based on theoretical models that do not include a stage of preference formation. Second, our results suggest that building social ties between people at different levels of the income distribution via social (communicative) fora is an important way to increase equality in groups where some members have a conflict of

interest between monetary self-interest and equality. Third, for group members who do not face a conflict of interest, our results suggest that group deliberation has an effect that cannot be explained by the building of psychological ties (social identity). Instead, social comparison and persuasion seem to explain the egalitarian shift of impartial members. If this holds more generally, then on-topic group deliberation has an important role to play in shaping the preferences of these group members that cannot be achieved by other social fora where people meet. Fourth, our results for the individual deliberation treatment suggest that there is a role for informed individual deliberation to play in shaping distributional preferences. Finally, from a more general point of view, the fact that distributional preferences change via individual and group deliberation makes it more difficult to use them as a "fixed" feature of individual preferences. This raises important questions for the elicitation of these preferences. While a discussion of this issue is beyond the scope of this essay, more attention is needed to model and understand how institutions like group deliberation form the (distributional) preferences that ultimately will feed into decision-making.

Some caution is needed when drawing inferences from our results. First, they are found in the highly stylized environment of the lab. The closest analog to our setting can be found in online deliberation platforms (e.g. message boards of news and social media websites/apps) that share many basic features of our experimental deliberation, but more caution is needed when using our results to inform debates about other forms of group deliberation. Second, the deliberating groups in our experiment are made up of five people. While this is larger than most economic experiments on communication and comparable to some teams, committees, and other small deliberating groups, caution is needed when using our results to draw conclusions about larger groups. Finally, we have implemented a deliberative procedure that — while retaining high experimental control — "promotes substantive, balanced, and civil discussion" (List 2018). Other deliberative procedures, such as those that allow people to communicate in silos (e.g. Brandts et al. 2022), may lead to different results.

Several extensions to our study are possible. For example, it would be interesting to replicate our design in a more complex (ethical) environment. The fact that the allocation of positions in the groups is purely random offers no room for broader justice considerations related to responsibility or merit, which may dominate the debate about redistribution in the real world. It is straightforward to extend our approach to contexts in which inequality is based on choice

luck (e.g. Cappelen et al. 2013) or on a mix of effort, productivity, and luck (e.g. Cappelen et al. 2010). One could also incorporate allocations that allow distinguishing between egalitarian and maximin preferences (as in Ueshima et al. 2021), or manipulate the structural features of the deliberative procedure (as, e.g., in Brandts et al. 2022). These extensions are left for future work.

Acknowledgments

We are grateful to Larbi Alaoui, Federica Alberti, Jordi Brandts, Alexander Cappelen, Francesco Cerigioni, Laurent Denant-Boemont, Nicolas Gravel, Olivier L'Haridon, Fabrice Le Lec, Stéphane Lemaire, David Masclet, Rosemarie Nagel, Antonio Penta, Rohini Somanathan, Joël van der Weele, and the seminar and workshop participants at Pompeu Fabra Economics, Portsmouth, Lille, Marseille, Besançon and Lyon, for valuable discussions, comments and suggestions. We are also grateful to Quentin Thévenet, Cecil Coskun, Oumayma Ben Khlifa, Thomas Cadet and Lamine N'Diaye for excellent research assistance. This paper forms part of the research project "ValFree" (The Value of Choice, grant No. ANR-16-CE41-0002-01) of the French National Agency for Research whose financial support is gratefully acknowledged. The experiment was approved by the GATE-LAB Ethical Board.

References

Ahn, T. K., B. Loukas, M. Batsaikhan, F. Campos-Ortiz, L. Putterman, and M. Sutter (2018) Trust and communication in a property rights dilemma.. Journal of Economic Behavior and Organization 149: 413–33.

Akerlof, G. A. (1980) A Theory of Social Custom, of Which Unemployment May Be One Consequence. The Quarterly Journal of Economics 94(4): 749–75.

Alesina, A. and P. Giuliano (2011) Preferences for redistribution. In: J. Benhabib, A. Bisin, and M. Jackson (eds) Handbook of Social Economics. 1A Elsevier, New York: 93–131.

Alesina, A., S. Stantcheva, and E. Teso (2018) Intergenerational mobility and support for redistribution. The American Economic Review 108: 521–54.

Almås, I., A. Cappelen, and B. Tungodden (2020) Cutthroat capitalism versus cuddly socialism: Are Americans more meritocratic and efficiency-seeking than Scandinavians?. Journal of Political Economy 128: 1753–88.

Almås, I., A. W. Cappelen, E. Sørensen, and B. Tungodden (2010) Fairness and the Development of Inequality Acceptance. Science 328: 1176–78.

Ambrus, A., B. Greiner, and P. Pathak (2015) How individual preferences are aggregated in groups: an experimental study. Journal of Public Economics 129: 1–13.

Andreoli, F. and J. Olivera (2020) Preferences for redistribution and exposure to tax-benefit schemes in Europe. European Journal of Political Economy 63(101880): 1–19.

Andreoni, J. and J. M. Rao (2011) The power of asking: How communication affects selfishness, empathy, and altruism. Journal of Public Economics 95: 513–20.

Balafoutas, L., R. Kerschbamer, M. Kocher, and M. Sutter (2014) Revealed distributional preferences: Individuals vs. teams. Journal of Economic Behavior and Organization 108: 319–30.

Balafoutas, L., M. Kocher, L. Putterman, and M. Sutter (2013) Equality, equity and incentives: An experiment. European Economic Review 60: 32–51.

Bernheim, B. D. (1994) A theory of conformity. Journal of Political Economy 102(5): 841–77.

Bonnet, I. and R. Zeckhauser (2004) Social Comparisons in Ultimatum Bargaining. Scandinavian Journal of Economics 106(3): 495–510.

Brandts, J., L. Gerhards, and L. Mechtenberg (2022) Deliberative structures and their impact on voting under economic conflict. Experimental Economics 25(2): 680–705.

Brandts, J., A. E. Giritligil, and R. A. Weber (2015) An experimental study of persuasion bias and social influence in networks. European Economic Review 80: 214–29.

Brown, R. (1974) Further comment on the risky shift. American Psychologist 29: 468–70.

——— (1986) R. Brown. The Free Press, New York2nd edition.

Burnstein, E. and A. Vinokur (1975) What a person thinks upon learning he has chosen differently from others: Nice evidence for the persuasive-arguments explanation of choice shifts.. Journal of Experimental Social Psychology 11: 412–26.

Burnstein, E., A. Vinokur, and Y. Trope (1973) Interpersonal comparisons versus persuasive argument: a more direct test of alternative explanations for group-induced shifts in individual choices. Journal of Experimental Social Psychology 9: 236–45.

Can, B. (2014) Weighted Distances Between Preferences. Journal of Mathematical Economics 51: 109–15.

Cappelen, A., J. Konow, E. Sorensen, and B. Tungodden (2013) Just luck: An experimental study of risk-taking and fairness.. The American Economic Review 103: 1389–413.

Cappelen, A., E. Sorensen, and B. Tungodden (2010) Responsibility for what? Fairness and individual responsibility. European Economic Review 54, 429-441. 54: 429–41.

Cappelen, A. W., A. D. Hole, E. Sørensen, and B. Tungodden (2007) The Pluralism of Fairness Ideals: An Experimental Approah. American Economic Review 97(3): 818–27.

Cason, T. and V.-L. Mui (1997) A laboratory study in group polarisation in the team dictator game. Economic Journal(107): 1465–83.

Cettolin, E. and A. Riedl (2017) Justice under uncertainty. Management Science 63(11): 3739–59.

Charness, G. and M. Rabin (2002) Understanding social preferences with simple tests. The quarterly journal of economics 117(3): 817–69.

Charness, G. and M. Sutter (2012) Groups make better self-interested decisions. Journal of Economic Perspectives 26(3): 157–76.

Chen, R. and Y. Chen (2011) The Potential of Social Identity for Equilibrium Selection. The American Economic Review 101(6): 2562–89.

Chen, Y., F. M. Harper, J. Konstan, and S. X. Li (2010) Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens. The American Economic Review 100(4): 1358–98.

Chen, Y. and S. X. Li (2009) Group Identity and Social Preferences. The American Economic Review 99(1): 437–57.

Christie, R. and F. L. Geis (1970) Studies in machiavellianism. Academic Press, New York, NY.

Costa-Font, J. and F. Cowell (2015) Social identity and redistributive preferences: A survey. Journal of Economic Surveys 29(2): 357–74.

Crawford, V. P. (1998) A Survey of Experiments on Communication via Cheap Talk. Journal of Economic Theory 78: 286–98.

Crawford, V. P. and J. Sobel (1982) Strategic information transmission. Econometrica 50(6): 1431–51.

Datta, S. and G. A. Satten (2008) A signed-rank test for clustered data. Biometrics 64(2): 501–7.

Dawes, R. M., A. J. Van De Kragt, and J. M. Orbell (1988) Not me or thee but we: The importance of group identity in eliciting cooperation in dilemma situations: Experimental manipulations. Acta Psychologica, 68(1-3), 83-97. 68(1-3): 83–97.

Durante, R., L. Putterman, and J. Van der Weele (2014) Preferences for redistribution and perception of fairness: An experimental study. Journal of the European Economic Association 12(4): 1059–86.

Dworkin, R. (2000) Sovereign Virtue. Harvard University Press, Cambridge MA.

Ellingsen, T. and M. Johannesson (2004) Promises, Threats and Fairness. The Economic Journal 114: 397–420.

Engelmann, D. and M. Strobel (2004) Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments. American Economic Review 94(4): 857–69.

Esteban, J-M and D. Ray (1994) On the Measurement of Polarization. Econometrica 62(4): 819–51.

Fishkin, J. S. (2018) Democracy When the People are Thinking: Revitalizing our Politics Through Public Deliberation. Oxford University Press, Oxford.

Fisman, R., K. Gladstone, I. Kuziemko, and S. Naidu (2020) Do Americans want to tax wealth? Evidence from online surveys. Journal of Public Economics 188(104207): 1–13.

Fisman, R., I. Kuziemko, and S. Vannutelli (2021) Distributional preferences in larger groups: Keeping up with the Joneses and keeping track of the tails. Journal of the European Economic Association 19(2): 1407–38.

Frey, B. S. and I. Bohnet (1997) Identification in democratic society. The Journal of Socio-Economics 26(1): 25–38.

Frey, B. S. and S. Meier (2004) Social Comparisons and Pro-Social Behavior: Testing "Conditional Cooperation" in a Field Experiment. The American Economic Review 94(5): 1717–22.

Gerlitz, Y. and J. Schupp (2005) Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP [Assessment of Big Five personality characteristics in the SOEP]. German Institute of Economic Research (Research Notes 4) Berlin: DIW.

Glazer, J. and A. Rubinstein (2004) On optimal rules of persuasion. Econometrica 72(6): 1715–36.

Goeree, J. K. and L. Yariv (2011) An Experimental Study of Collective Deliberation. Econometrica 79(3): 893–921.

Heap, S. P. H., K. Tsutsui, and D. J. Zizzo (2020) Vote and voice: An experiment on the effects of inclusive governance rules. Social Choice and Welfare 54(1): 111–39.

Kamenica, E. and M. Gentzkow (2011) Bayesian persuasion. The American Economic Review 101(6): 2590–615.

Kemeny, J. (1959) Mathematics without Numbers. Daedalus 88(4): 577–91.

Kittel, B., W. Luhan, and R. Morton (2014) Communication and voting in multi-party elections: An experimental study. The Economic Journal 124(574): 196–225.

Konow, J. (2000) Fair shares: Accountability and cognitive dissonance in allocation decisions. The American Economic Review 90(4): 1072–91.

Krupka, E. L. and R. A. Weber (2009) The focusing and informational effects of norms on pro-social behavior. Journal of Economic Psychology 30: 307–20.

——— (2013) Identifying social norms using coordination games: Why does dictator game sharing vary?. Journal of the European Economic Association 11(3): 495–524.

Kuziemko, I., M. Norton, E. Saez, and S. Stantcheva (2015) How elastic are preferences for redistribution? Evidence from randomized survey experiments. The American Economic Review 105: 1478–1508.

List, C. (2018) Democratic Deliberation and Social Choice: A Review. In: A. Bächtiger, J. S. Dryzek, J. Mansbridge, and M. Warren (eds) The Oxford Handbook of Deliberative Democracy. Oxford University Press, Oxford.

List, C., R. Luskin, J. S. Fishkin, and I. McLean (2013) Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls. Journal of Politics 75: 80–95.

List, J. A. (2007) On the interpretation of giving in dictator games. Journal of Political economy 115(3): 482–93.

Luhan, W. J., M. G. Kocher, and M. Sutter (2009) Group polarization in the team dictator game reconsidered. Experimental Economics 12: 26–41.

Mengel, F. and E. Weidenholzer (2022) Preferences for redistribution. Journal of Economic Surveys: 1–18.

Myers, D. G. and H. Lamn (1976) The Group Polarization Phenomenon. Psychological Bulletin 83(4): 602–27.

Penczynski, S. P. (2016) Persuasion: An experimental study of team decision making. Journal of Economic Psychology 56: 244–61.

Pruitt, D. G. (1971a) Choice shifts in groups discussion: An introductory review. Journal of Personality and Social Psychology 20: 339–60.

——— (1971b) Conclusions: Toward an understanding of choice shifts in group discussion. Journal of Personality and Social Psychology 20: 495–510.

Schokkaert, E. and B. Tarroux (2021) Empirical research on ethical preferences: How popular is prioritarianism?. In: M. Adler and O. Norheim (eds) Prioritarianism in Practice. Cambridge University Press, Cambridge.

Sunstein, C. R. (2002) The Law of Group Polarization. The Journal of Political Philosophy 10(2): 175–95.

Sutter, M., M. Kocher, and M. Praxmarer (2020) Team Decision-Making. In: K. F. Zimmermann (ed) Handbook of Labor, Human Resources and Population Economics. Springer, Cham, : 1–25.

Tajfel, H. and J. C. Turner (1979) An Integrative Theory of Intergroup Conflict. In: W. G. Austin and S. Worchel (eds) The Social Psychology of Intergroup Relations. Wadsworth, Monterey, CA.

Ueshima, A., H. Mercier, and T. Kameda (2021) Social deliberation systematically shifts resource allocation decisions by focusing on the fate of the least well-off. Journal of Experimental Social Psychology 92(104067).

Weinzierl, M. (2017) Popular acceptance of inequality due to innate brute luck and support for classical benefit-based taxation. Journal of Public Economics 155: 54–63.

Yang, Z., X. Sun, and J. W. Hardin (2010) A note on the tests for clustered matched-pair binary data.. Biometrical Journal 52(5): 638–52.

Zizzo, D. J. (2010) Experimenter demand effects in economic experiments. Experimental Economics 13(1): 75–98.

ONLINE APPENDIX

How group deliberation affects individual distributional preferences: An experimental study

João V. Ferreira, Erik Schokkaert, Benoît Tarroux

Table of Contents

A. Appendix to the main results

For the decomposition, we start from the following regression specification[11]:

[11] We drop the subscript *i* for notational convenience.

$$\alpha = \beta_0 + \sum_{g=2}^{3} \beta_{1g} T_1 p_g + \sum_{g=1}^{3} \beta_{2g} T_2 p_g + \beta_3 S + \varepsilon \tag{5}$$

where T_1 and T_2 are binary variables for treatment, p_g is a binary variable for period $g = \{1,2,3\}$, S are dummy variables for sessions and ε is a zero-expected noise. If we use the (discrete) choice of an allocation as our measure of the degree of egalitarianism, we estimate an ordered logit model and expression (5) should be interpreted as the specification of the latent variable underlying the discrete choice process. If we measure α by the Kemeny distance, we use an OLS regression.[12]

It is also worth noting that we impose two constraints in our main decomposition. First, note that β_0 measures the value of the Kemeny distance at period 1 for the subjects in T1. The analogous measure for the subjects in T2 is $\beta_0 + \beta_{21}$, and if our randomization has worked β_{21} should be equal to zero. The hypothesis that β_{21} is different from zero is indeed rejected in our results. We therefore impose $\beta_{21} = 0$ in our estimation. Second, our conceptual approach also assumes that $\beta_{22} = \beta_{23}$, i.e., that the mere fact of switching groups in T2 does not change the degree of egalitarianism obtained from individual deliberation with social comparison. Both coefficients are estimated with a large margin of error, but the hypothesis of equality cannot be rejected (Wald tests: $p = 0.805$ for column 1 and $p = 0.455$ for column 3). Given that $\beta_{22} = \beta_{23} = \widetilde{\beta}$ is not rejected, we also impose it in our estimation. The constrained estimation results are shown in the second and the fourth column of Table A1. Imposing these constraints has the advantage that the effect of each mechanism, as described in expression (3), can be unambiguously identified using the estimated coefficients of expression (5):

$$\begin{aligned} \text{Total effect}: &\quad \beta_{12} \\ \text{Social identity}: &\quad \beta_{12} - \beta_{13} \\ \text{Social comparison}: &\quad \widetilde{\beta} \\ \text{Persuasion}: &\quad \beta_{13} - \widetilde{\beta} \end{aligned} \tag{6}$$

[12] Tobit estimates give similar results. These results are available upon request.

	(1)	(2)	(3)	(4)
	Preferred allocation		Kemeny distance	
	Unconstrained	Constrained	Unconstrained	Constrained
T1 × Period 1	(ref)	(ref)	(ref)	(ref)
T1 × Period 2	0.880***	0.904***	0.189***	0.193***
	(0.137)	(0.144)	(0.027)	(0.029)
T1 × Period 3	0.396***	0.421***	0.099***	0.102***
	(0.098)	(0.117)	(0.020)	(0.025)
T2 × Period 1	-0.191	0	-0.023	0
	(0.535)		(0.123)	
T2 × Period 2	-0.046	0.091	0.029	0.037
	(0.551)	(0.233)	(0.123)	(0.051)
T2 × Period 3	-0.021	0	0.017	0
	(0.542)		(0.124)	
Observations	720	720	720	720
R^2			0.053	0.053
Log Likelihood.	-806.505	-806.622		

Notes: This table reports the regression estimates of eq. 5 without any constraints (columns 1 and 3) and assuming $\beta_{12} = 0$ and $\beta_{22} = \beta_{23}$ (columns 2 and 4). *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level, based on an ordered logit model for the preferred allocation in columns 1 and 2 and an OLS for the K-distance in columns 3 and 4, with clustered standard errors per subject and session dummies.

Table A1: Regression estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Rich subjects		Median subjects		Poor subjects	
	Pref. alloc.	K-distance	Pref. alloc.	K-distance	Pref. alloc.	K-distance
T1× Period 1	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
T1× Period 2	2.000***	0.311***	1.594***	0.217**	0.287	0.062**
	(0.341)	(0.049)	(0.454)	(0.059)	(0.427)	(0.028)
T1× Period 3	0.460	0.083**	1.165***	0.197***	0.486	0.073***
	(0.329)	(0.035)	(0.389)	(0.053)	(0.403)	(0.024)
T2× Period 2 / 3	-0.379	0.015	0.741	0.062	1.168	0.0471**
	(0.635)	(0.049)	(0.572)	(0.077)	(0.735)	(0.021)
Observations	288	288	144	144	288	288
R^2		0.230		0.244		0.109
Log Likelihood	-241.668		-152.234		-133.704	

Notes: This table reports the regression estimates of eq. 5 assuming $\beta_{12} = 0$ and $\beta_{22} = \beta_{23}$ and per rank groups. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level, based on an ordered logit model for the preferred allocation in columns 1, 3, and 5 and an OLS for the K-distance in columns 2, 4 and 6, with clustered standard errors per subject and session dummies.

Table A2: Regression estimates per rank groups

Difference between periods in the fraction of subjects preferring:							
	a1	a2	a3	a4	a5	a6	a7
p1=p2	<0.001	0.006	0.659	1.000	1.000	0.126	<0.001
p1=p3	0.015	0.033	1.000	0.120	0.779	0.062	0.023
p2=p3	0.004	0.184	0.587	0.168	0.592	0.780	0.042

Notes: This table reports *p*-values for the extended Stuart-Maxwell tests of homogeneity of marginal distributions with clustered matched data and clustering at the group level for p1=p2 and at the session level for p1=p3 and p2=p3 (Yang et al. 2010).

Table A3: Differences in preferred allocations between periods

B. Results for T2

In this appendix, we present the results on the effect of individual deliberation with social comparison on distributional preferences (treatment 2). Figure A1 reports the distribution of preferred allocations over the three periods and Table A4 reports the changes in Kemeny distance in this treatment. Comparing periods 1 and 2, we observe that individual deliberation with social comparison has an egalitarian effect. However, this effect is much smaller in magnitude than in T1 and is only statistically significant for the Kemeny distance ($p = 0.048$). In fact, the effect of deliberation (period 2 - period 1) in T1 is statistically significantly larger than in T2 ($p = 0.007$ and $p = 0.020$, Wald test based on columns 2 and 4 of Table A3 respectively). From Figure A1 and Table A4, we can also observe that — as predicted by SCT — there are no statistically significant differences between periods 2 and 3. When looking at the results per rank groups (Table A5), we observe that the effect of individual deliberation with social comparison is only statistically significant for median (impartial) members.

C. The degree of egalitarianism in a utility model

	Mean	% with K-distance			Signed-rank tests	
		$= 0$	$\in]0,1[$	$= 1$	= period 2	= period 3
Period 1	0.52	33.85	23.08	43.08	$p = 0.048$	$p = 0.111$
Period 2	0.57	35.38	16.92	47.69		$p = 0.308$
Period 3	0.56	36.92	15.38	47.69		

Notes: Signed-rank test for clustered data, clustering at the group level for p1=p2 and at the session level for p2=p3 and p1=p3.

Table A4: Kemeny distances in T2

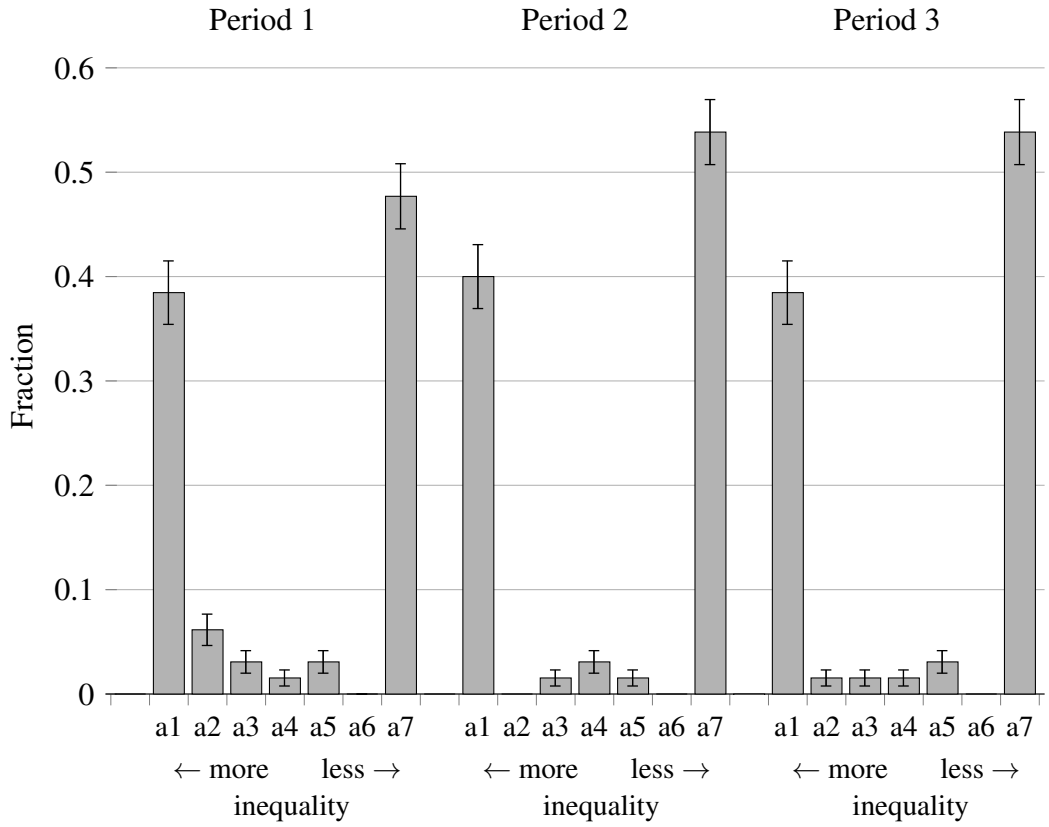

Figure A1: Distribution of preferred allocations in T2 across periods

Notes: The main bars display percentages of subjects with preferred allocations a1 to a7 per treatment. The error bars display one standard error of the mean, with values based on logit estimates with dummies for allocations for each period.

	Mean	% with K-distance			Signed-rank tests	
		$=0$	$\in]0,1[$	$=1$	= period 2	= period 3
Rich subjects						
Period 1	0.08	76.92	19.23	3.85	$p=0.900$	$p=0.386$
Period 2	0.12	80.77	15.38	3.85		$p=0.155$
Period 3	0.06	84.62	15.38	0.00		
Median subjects						
Period 1	0.56	7.69	61.54	30.77	$p=0.045$	$p=0.152$
Period 2	0.71	7.69	46.15	46.15		$p=0.815$
Period 3	0.74	7.69	46.15	46.15		
Poor subjects						
Period 1	0.94	3.85	7.69	88.46	$p=0.150$	$p=0.158$
Period 2	0.96	3.85	3.85	92.31		$p=0.311$
Period 3	0.96	3.85	0.00	96.15		

Notes: Signed-rank test for clustered data, clustering at the group level for p1=p2 and at the session level for p2=p3 and p1=p3.

Table A5: Kemeny distances across periods per rank groups in T2

We complete the results based on the choice of the preferred allocation and the Kemeny distance with an analysis by a random utility model. We assume a simplified version of the Charness and Rabin (2002) model in which the individual arbitrates between her own payoff and the minimal payoff in the group. This simplification is driven by two features of our setting: allocations are set up such that (i) minimal payoff is perfectly aligned with equality, and (ii) minimal payoff/equality run in almost linear opposition to efficiency as total payoff. Therefore, it would not be possible to estimate (with confidence) a model that has a trade-off between own payoff, minimal payoff, equality, and efficiency. Since own payoff and minimal payoff are equal for the poorest member, we exclude these subjects from the analysis.

Denote by $x_k = (x_{1k}, ..., x_{5k})$ an allocation with $x_{1k} \leq ... \leq x_{5k}$. Assume that the utility of i from allocation k at period t, denoted u_{ik}^t, depends on his own gain x_{ik} and the minimal gain x_{1k} of the group:

$$u_{ik}^t = v_{ik}^t + \varepsilon_{ik}^t = \alpha_t x_{ik} + \beta_t x_{1k} + \varepsilon_{ik}^t \tag{7}$$

where the ε_{ik}^t's are independently and identically distributed with an extreme-value distribution. Assuming $\alpha_t + \beta_t = 1$ yields:

$$u_{ik}^t = (1-\beta_t)x_{ik} + \beta_t x_{1k} + \varepsilon_{ik}^t = \beta_t(x_{1k} - x_{ik}) + x_{ik} + \varepsilon_{ik}^t \tag{8}$$

In our experimental setup, subjects rank 7 allocations and the parameters of equation (8) can be estimated with a rank-ordered logistic model. We assume that β can be different in the three periods and the two treatments. The utility can then be expressed as follows:

$$u_{ig}^t = T_1 \times \sum_{g=1}^{3} p_g\big(\beta_{1g}(x_1 - x_i)\big) + T_2 \times \sum_{g=1}^{3} t_g\big(\beta_{2g}(x_1 - x_i)\big) + x_i + \varepsilon_{ik}^t$$

where p_g is a binary variable for period $g = \{1,2,3\}$, and T_1 and T_2 are binary variables indicating the treatment.

As in Section 5.2, we impose the constraints $\beta_{11} = \beta_{21} = \beta_{\cdot 1}$ and $\beta_{22} = \beta_{23} = \widetilde{\beta}$. Table A7 presents the decomposition analysis for the whole sample and per rank groups. The results confirm the findings described in the main text. The minimal income in the group has a larger weight in the individual utility function after than before collective deliberation. Moreover, β_{12} is not only statistically significantly higher than β_{13} ($p < 0.001$, Wald test), but also statistically

	All subjects	Rich subjects	Median subjects	2nd Poorest subjects
$\beta_{.1}$ (Period 1)	0.086*	0.045	0.132**	0.363
	(0.050)	(0.074)	(0.070)	(0.377)
β_{12} (Period 2)	0.470***	0.463***	0.519***	1.590***
	(0.051)	(0.044)	(0.208)	(0.453)
β_{13} (Period 3)	0.268***	0.215***	0.480***	2.130**
	(0.059)	(0.078)	(0.164)	(0.705)
$\widetilde{\beta}$ (Period 2/3)	0.071	-0.102	0.258*	0.745
	(0.053)	(0.101)	(0.083)	(1.241)
Observations	5040	2016	1008	1008
Pseudo-R^2	0.158	0.233	0.177	0.133
Log Likelihood	-3967.204	-1762.620	-1010.382	-475.691

Notes: *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level, based on a rank-ordered logit model with clustered standard errors per subject.

Table A6: Rank-ordered logit estimates

	Total	Identity	Social Comp.	Persuasion	SC+P
All subjects					
β **(utility function)**	0.385	0.203	-0.015	0.197	0.182
Wald tests (p-values)	0.000	0.000	0.840	0.015	0.000
Rich subjects					
β **(utility function)**	0.418	0.248	-0.147	0.317	0.170
Wald tests (p-values)	0.000	0.000	0.407	0.074	0.011
Median subjects					
β **(utility function)**	0.387	0.039	0.126	0.222	0.348
Wald tests (p-values)	0.004	0.564	0.370	0.265	0.003
2nd Poorest subjects					
β **(utility function)**	1.227	-0.541	0.382	1.386	1.768
Wald tests (p-values)	0.077	0.552	0.753	0.390	0.008

Notes: This table reports the decomposition results using the regression estimates reported in Table A6. SC+P stands for the joint effect of social comparison and persuasion.

Table A7: Decomposition results

significantly higher than $\widetilde{\beta}$ ($p < 0.001$, Wald test). As a result, the decomposition suggests that social identity (measured by $\beta_{12} - \beta_{13}$) and persuasion (measured by $\beta_{13} - \widetilde{\beta}$) are both statistically significant mechanisms behind the observed egalitarian shift.

When looking at the two richest group members, their degree of egalitarianism is statistically significantly larger in period 2 than in periods 1 and 3, and in period 3 than in period 1 ($p < 0.001$ for all tests, Wald tests). Again, the decomposition suggests that the changes are driven by social identity and persuasion (although the latter effect is only significant at 10%). As with the other proxies, the median members are more egalitarian after group deliberation and more egalitarian in period 3 than in period 1 ($p < 0.01$ for both cases, Wald tests), but we observe no change between the second and third periods: the difference between β_{12} and β_{13} is not statistically significant ($p = 0.564$, Wald test). This implies that social identity does not play a role for them and their egalitarian shift is only driven by the joint effect of persuasion and social comparison. Due to the high correlation between the minimal income and the income of the second poorest member, the parameters of their utility function cannot be reliably estimated.

D. Herfindahl index of group polarization

In this appendix, we present the results for an alternative measure of group polarization named the normalized Herfindahl index — which using the preferred allocation as the measure for the degree of egalitarianism — is defined as follows:

$$H^* = \frac{H - \underline{H}}{\overline{H} - \underline{H}} \text{ where } H = \sum_{l=1}^{7} (\pi_{l,Td}^t)^2. \tag{9}$$

and $\pi_{l,Td}^t$ is the proportion of subjects with a degree of egalitarianism $\overline{\alpha}_l$ at period t in treatment d. If all N subjects have a different degree of egalitarianism, H reaches its minimal value $\underline{H} = 1/N$. If there is full consensus, in that all subjects have the same degree of egalitarianism, H reaches its maximal value $\overline{H} = 1$. The Herfindahl index H can be seen as measuring the "closeness to consensus". Here we use the normalized version of the Herfindahl index, denoted H^*, which will take values between 0 (when all disagree) and 1 (when there is consensus).

The results for T1 and the preferred allocation are shown in Table A8. In a nutshell, these results are equivalent to the ones found for the other polarization measures. The normalized

	Periods			Signed-rank tests		
	p1	p2	p3	p1 = p2	p1 = p3	p2 = p3
Norm. Herfindhahl index	0.30	0.55	0.35	$p < 0.001$	$p < 0.001$	$p < 0.001$

Notes: Signed-rank test for clustered data, clustering at the group level for p1=p2 and at the session level for p2=p3 and p1=p3.

Table A8: Polarization within groups (normalized Herfindahl index)

Herfindahl strongly and significantly increases in period 2 (i.e., preferences get closer to consensus after group deliberation), it then decreases in period 3, while remaining statistically significantly higher than in period 1.

E. Chat content analysis

The analysis of the content of the chat scripts can bring additional insights into the psychological mechanisms at work. In this section, we (i) describe the methodology we use to analyze the chat scripts, (ii) show the agreement between different coders and compare it against a random benchmark, (iii) describe the content of the chats, and finally (iv) analyze the effect of the chat content on subjects' distributional preferences.

Methodology

Three research assistants read all the chats and were asked to classify and code the messages according to 24 categories. We aggregate some of these categories to arrive at a classification with a sufficiently large proportion of messages in each category. A first set of categories concerns (reasoned) arguments in favor or against the different allocations. Among these, the category "pro-social arguments" gathers messages mentioning equality or inequality (e.g., 'inequality is higher in allocation 1'), fairness (e.g., 'allocation 7 is the fairest'), maximin (e.g. 'with allocation 1 members D and E will depart without much'), or the general interest (e.g., 'allocation 5 is the best for the group'). An "efficiency argument" evokes a defense of a ranking based on the amount of total income (e.g. 'the total payoff is higher in allocation 1'). The category "libertarian arguments" gathers messages stating that one should choose whatever they please (e.g. 'everyone should choose the best for themselves'), and an "appeal to sympathy" consists in telling others to put themselves in their shoes (e.g., 'imagine being in my position'). A "compromise" is a proposal to do something if others do something in return, or a plea for others to choose an allocation that is a compromise. A second set of categories correspond to

	# of messages per group		# of groups for which # of messages > 0
	Average	σ	
Total	52.94	22.87	35 out of 35
Arguments			
Pro-social argument	5.54	3.68	34
Efficiency argument	0.51	0.92	12
Libertarian argument	0.26	0.61	7
Appeal to sympathy	0.34	0.84	7
Compromise	0.31	0.76	7
Simple responses and propositions			
Agreement with other(s)	7.23	3.87	35
Disagreement with other(s)	1.49	1.87	21
Ask opinion	4.80	2.82	35
Simple proposition	4.26	4.23	31
Irrelevant			
Out-of-topic	12.94	15.18	27
Greetings and goodbyes	3.89	2.45	28
Ambiguous	0.66	1.25	13
Confusion	0.66	0.97	15
Not-classified messages	7.94	8.12	29
# of groups reaching an agreement: 31 out of 35			
(1 group on allocation 1, 7 groups on allocations 2-6, and 23 groups on allocation 7)			

Table A9: Main categories and content of the group chats

simple responses and propositions, such as those in which subjects express their agreement or disagreement with others (e.g., 'that's right'), ask the opinion of others, or make a proposition without mentioning an underlying reason for it. A third set contains greetings, goodbyes, and out-of-topic messages. Finally, the research assistants reported whether the groups reached a (non-binding) agreement or not, and, if any, on which allocation the members agreed upon.[13]

We assigned a message to a specific category if at least 2 out of the 3 coders agreed. The three coders only (fully) disagreed on the content of 7.94 messages per group (15% of the total number of messages per group), and these were not classified.

[13]We also asked the research assistants to evaluate the emotional content of each message: friendly/positive, neutral or unfriendly/negative. With an average of 52.94 messages per group, the overwhelming majority of messages were classified as neutral (46.83 on average per group), a few positive (5.77), and almost none negative (0.34). Including this data in our analysis does not change our results. For the sake of concision, we omit this from the paper, and the results are available upon request.

Agreement between coders

To further measure the degree of agreement between two coders, we determine for each message posted by a subject the difference in evaluation between coders. For each message m, coder i is asked to assign 1 if a message belongs to a category k and 0 if it does not. Denote v_{imk} the binary variable taking the value 1 if $i \in \{1,2,3\}$ classifies $m \in \{1,...,M\}$ in category $k \in \{1,...,K\}$, and 0 otherwise. The degree of agreement of two coders about a message m is then the fraction of categories in which both coders classify a message. That is, the degree of agreement, denoted by a_{ijm}, might be written as:

$$a_{ijm} = \frac{1}{K} \sum_{k} \mathrm{I}(v_{imk} = v_{jmk})$$

where $\mathrm{I}(z) = 1$ if condition z is true and 0 otherwise. The overall degree of agreement is simply the average value of a_{ijm}:

$$a_{ij} = \frac{1}{M} \sum_{m} a_{ijm}$$

As v_{imk} is very likely to take the value of 0, the value taken by a_{ijm} or a_{ij} is likely to be high. To avoid such a problem, we use a close measure of agreement:

$$\widetilde{a}_{ijm} = \frac{\sum_k \mathrm{I}(v_{imk} = v_{jmk})}{\sum_k \mathrm{I}(v_{imk} \neq 0 \ \& \ v_{jmk} \neq 0)} \quad \text{and} \quad \widetilde{a}_{ij} = \frac{1}{M} \sum_{m} \widetilde{a}_{ijm}$$

To test the accuracy of the evaluation given by our coders, we compare their classification with that of 100 'bot' coders. These bots randomly classify each message but select the same expected number of categories as our human coders (e.g., if our coders have selected 2 categories on average for a message m, the bots do the same). We then determine the overall agreement between each coder i and bots b by calculating a_{ib} and $\widetilde{a}_{ib}$.

Table A10 reports the degree of agreement between coders and between coders and bots. The agreement between coders is satisfactory: it is very high when considering the categories chosen *and* the ones not chosen for each message (a_{ij}), and coders coincide between 40 and 60% of the time when only taking into account the categories chosen for each message ($\widetilde{a}_{ij}$). In addition, the evaluations made by the coders are considerably closer to each other than to the bots. This is particularly visible for $\widetilde{a}_{ij}$, where the average agreement between each coder and

	a_{ij}	$\widetilde{a}_{ij}$
Coder 1 vs Coder 2	0.94	0.40
Coder 1 vs Coder 3	0.96	0.60
Coder 2 vs Coder 3	0.95	0.41
Coder 1 vs Bots	0.89	0.05
Coder 2 vs Bots	0.90	0.05
Coder 3 vs Bots	0.90	0.05

Table A10: Coders' reliability

the bots is about 5%. In other words, the coders' agreement is not random.

Chat content

Table A9 provides an overview of the chat content in relation to the categories described above. Groups exchanged on average more than 50 messages in the 10 minutes they had for deliberation. The shortest and longest discussions were of 12 and 97 messages respectively.[14] In terms of content, the chat scripts show that all groups engaged in discussion about the allocations. Groups shared on average more than 6 (reasoned) *arguments* in favor or against different allocations (see Table A9). As it can be seen from the table, pro-social arguments clearly dominate. In fact, there is only one group in which this type of argument is not invoked. Interestingly, a significant part of the messages can be classified as on-topic *simple responses and propositions*, suggesting that the deliberation is active and interactive: we find that subjects often express their agreement or disagreement with others, ask others about their opinion, or make a proposition without giving an explicit justification for it. The number of irrelevant messages still represents a non-negligible part of the discussion. This is not surprising, as the discussion was not restricted and subjects had enough time to exchange out of topic. Finally, we observe that 31 groups reached an agreement about the best allocation, and 23 of them converged towards the most egalitarian allocation 7.

The effect of the chat content on distributional preferences

Table A11 reports the results of an econometric analysis of the effects of the content of the chats on subjects' degree of egalitarianism. The main findings can be summarized as follows.

[14]The average number of words in a message was 6.66 with a standard deviation of 5.83. The shortest messages contained one word and the longest one had 49 words.

First, the number of (reasoned) arguments has a statistically significant effect on subjects' degree of egalitarianism. This is perfectly in line with the PAT hypothesis that people are not only influenced by the persuasiveness of the arguments, but also by the sheer number of arguments put forward in favor or against a given view. As expected, pro-social arguments have a positive effect on the propensity of subjects to adopt egalitarian preferences. This effect is statistically significant and particularly prominent given its magnitude and the average number of pro-social arguments exchanged per group. Note that this effect is (slightly) concave. Also as expected, the effect of efficiency arguments goes in the opposite direction. This effect is also statistically significant, in particular for the Kemeny distance. Libertarian arguments have an effect on the preferred allocation but fail the significance test for the Kemeny distance, while appeals to sympathy seem to have no effect on subjects' distributional preferences. Arguments for compromise have a large negative effect on subjects' degree of egalitarianism, suggesting that compromise is sought by rich members. Finally, we observe that the number of irrelevant out-of-topic messages and the length of the discussion do not impact subjects' degree of egalitarianism.

	(1) Preferred allocation	(2) K-distance
Preferred in period 1		
Allocation 1	(ref)	
Allocations 2 to 6	0.100	
	(0.132)	
Allocation 7	-0.640***	
	(0.071)	
K-distance in period 1		-0.431***
		(0.049)
# Pro-social	0.106***	0.054***
	(0.024)	(0.019)
(# Pro-social)^2	-0.007***	-0.004***
	(0.002)	(0.001)
# Efficiency	-0.048*	-0.045***
	(0.024)	(0.016)
# Libertarian	0.101***	0.040
	(0.033)	(0.025)
# Appeal to sympathy	-0.019	0.011
	(0.020)	(0.016)
# Compromise	-0.086***	-0.095***
	(0.025)	(0.021)
Agreement allocation 1-6	(ref)	(ref)
No agreement	0.114**	0.013
	(0.055)	(0.050)
Agreement allocation 7	0.320***	0.162***
	(0.069)	(0.050)
# Irrelevant	0.005**	0.001
	(0.002)	(0.002)
Total # of messages	-0.003	0.000
	(0.002)	(0.001)
Constant	0.110	0.187**
	(0.107)	(0.089)
Observations	175	175
R^2	0.437	0.387
Log Likelihood	-93.672	

Notes: *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level, based on an ordered logit model for the preferred allocation in column 1 and an OLS for the K-distance in column 2 with clustered standard errors per group. The dependent variable in column 1 is equal to -1 if the subject prefers a less egalitarian allocation in period 2; 0 if she reports the same preferred allocation in both periods; and 1 if she prefers a more egalitarian allocation in period 2. The dependent variable in column 2 is the change in the K-distance between periods 1 and 2.

Table A11: Chat content and changes in the degree of egalitarianism between periods 1 and 2

F. Conformity

In this appendix, we check if conformity can be an explanatory mechanism of the preference change we observe after group deliberation. Recall that according to this theory, subjects will converge towards the numerical majority/median of the group. To distinguish between conformity and social comparison, we can look at the groups where the numerical majority/median is "not egalitarian" and check if these groups converge on this non-egalitarian position.

For T1, contrary to conformity's prediction, we observe a systemic convergence towards the most egalitarian allocation no matter the group's median preferred allocation in the first period (see Table A12).[15] For example, the shift in Kemeny distance is always positive with one exception for which the shift is practically null (see the last column). The relatively smaller egalitarian shift when allocation 1 was the group's median preferred allocation could be explained by the effect of conformity, but also by the effect of persuasion in these groups being skewed in this direction.

For a further test of the effect of conformity, we can look at the results for T2. In this case, conformity and social comparison can be at play, but other mechanisms such as persuasion and social identity are excluded. The two remaining theories have opposite predictions when the median is not egalitarian: conformity predicts convergence towards allocation 1 (if this is the group's median/numerical majority), while social comparison still predicts convergence towards allocation 7 as long as allocation 7 is the preferred allocation for one subject in the group. Let us look at the 6 groups (30 subjects) for which allocations 1, 2 and 3 — the three least egalitarian allocations — are the median preferred allocation in period 1 (see Table A13). For all these groups, there is at least one subject preferring allocation 7 as required by the social comparison mechanism. As it can be seen from Table A13, all groups either converge towards the most egalitarian allocation or have no discernible change in preferences no matter the group's median preferred allocation in the first period. This result aligns with the prediction of social comparison. We also do not find evidence that egalitarian choices increase as subjects observed more egalitarian choices on the part of others. It follows that if conformity plays a role in our setting, it does not seem to be strong and it cannot explain our results.

[15] We use the median instead of the numerical majority because the median allows us to classify all groups and because for the most "extreme" allocations the median and numerical majority are equivalent or similar. For example, the median at allocation 1 is equivalent to the numerical majority at allocation 1, and the median at

Group's median pref. allocation in period 1	Subjects' preferences in period 2							
	N	Pref. allocation		Δ Pref. allocation			K-distance	Δ K-distance
		a1	a7	Less equal	Same	More equal		
All subjects								
1	40	17	17	4	26	10	0.519	0.141
2	15	3	8	3	6	6	0.638	0.206
3	5	1	4	0	3	2	0.800	0.381
4	25	5	9	3	14	8	0.678	0.189
5	15	2	12	0	9	6	0.803	0.216
6	0							
7	75	9	64	0	54	21	0.810	0.194
All	175	37	114	10	112	53	0.709	0.189
Subjects whose preferred allocation is different from the group median one								
1	16	1	12	4	10	2	0.893	0.033
2	12	2	7	2	6	4	0.687	0.183
3	4	1	3	0	3	1	0.750	0.250
4	20	5	8	3	12	5	0.662	0.143
5	12	2	9	0	9	3	0.814	0.242
7	30	9	19	0	9	21	0.554	0.487
All	94	20	58	9	49	36	0.693	0.256
Subjects whose preferred allocation is equal to the group median one								
1	24	16	5	0	16	8	0.270	0.212
2	3	1	1	1	0	2	0.444	0.302
3	1	0	1	0	0	1	1.000	0.905
4	5	0	1	0	2	3	0.743	0.371
5	3	0	3	0	0	3	0.762	0.111
7	45	0	45	0	45	0	0.981	-0.001
All	81	17	56	1	63	17	0.728	0.112

Table A12: Preferred allocation and Kemeny distances in period 2 and group median choices in period 1 (T1)

Group's median preferred allocation in period 1	Subjects' preferences in period 2							
	N	Pref. allocation		Δ Pref. allocation			K-distance	Δ K-distance
		a1	a7	Less equal	Same	More equal		
All subjects								
1	10	5	5	0	9	1	0.467	0.100
2	10	4	4	0	8	2	0.538	0.100
3	10	4	6	2	6	2	0.586	0.086
4	5	2	3	0	4	1	0.600	0.086
5	5	1	3	0	3	2	0.695	0.124
6	0							
7	25	10	14	3	20	2	0.589	-0.021
All	65	26	35	5	50	10	0.571	0.052
Subjects whose preferred allocation is different from the group median one								
1	4	0	4	0	4	0	0.917	0.000
2	8	4	4	0	8	0	0.548	0.042
3	8	3	5	1	6	1	0.625	0.095
4	4	2	2	0	4	0	0.500	0.000
5	4	1	2	0	3	1	0.702	0.119
7	10	8	2	0	8	2	0.152	0.043
All	38	18	19	1	33	4	0.510	0.053
Subjects whose preferred allocation is equal to the group median one								
1	6	5	1	0	5	1	0.167	0.167
2	2	0	0	0	0	2	0.500	0.333
3	2	1	1	1	0	1	0.429	0.048
4	1	0	1	0	0	1	1.000	0.429
5	1	0	1	0	0	1	0.667	0.143
7	15	2	12	3	12	0	0.879	-0.064
All	27	8	16	4	17	6	0.656	0.051

Table A13: Preferred allocation and Kemeny distances in period 2 and group median choices in period 1 (T2)

G. Subjects socio-demographic characteristics

	All	T1	T2	*p*-value
Age (mean)	23.32	23.4	23.0	0.706 (Student)
Age (median)	22	22	22	0.887 (Mann-Whitney)
Female (%)	48.75	49.14	47.69	0.842 (Chi-Square)
Student (%)	90.42	89.14	93.85	0.271 (Chi-Square)
Field of study (%)				
Economics & Management	67.92	66.29	72.31	0.374 (Chi-Square)
Engineering	20.42	21.71	16.92	0.413 (Chi-Square)
Other	11.67	12.00	10.77	0.792 (Chi-Square)
Perceived social category (%)				
Rich	15.42	17.71	9.23	0.106 (Chi-Square)
Upper middle class	35.00	32.00	43.08	0.110 (Chi-Square)
Middle class	30.42	30.29	30.77	0.942 (Chi-Square)
Bottom middle class	12.92	13.14	12.31	0.864 (Chi-Square)
Poor	6.25	6.86	4.62	0.524 (Chi-Square)

Table A14: Characteristics of the subject's pool

H. The effect of personality traits and perceived deliberation quality

In this appendix, we estimate two regression models to evaluate the effects of socio-demographics, personality traits, and the perceived quality of the deliberation on changes to subjects' degree of egalitarianism before and after group deliberation. Table A15 reports the regression estimates using the pool of all subjects and the subsample of the rich and median members.[16] Somewhat surprisingly, we find no statistically significant effect of the personality traits such as the five categories of the Big Five questionnaire (Negative emotions, Extraversion, Open mindedness, Agreeableness, Conscientiousness), degree of Machiavelism, and social open mindedness (see Appendix K for the questionnaires underlying these measures). On the other hand, the perceived quality of the deliberation seems to have a positive effect on the likelihood of rich and median subjects becoming more egalitarian after group deliberation (measured 1 to 7 in increasing quality based on 6 questions about the deliberation phase; see Appendix K for underlying questions).

allocation 2 demands that there are at least 3 out of 5 group members with preferred allocations 1 and 2.

[16] We have 160 (T1) and 60 (T2) subjects instead of 175 and 65 subjects because the removed subjects did not respond to the short online survey with the three psychometric questionnaires that took place online before the lab session.

	(1)	(2)	(3)	(4)
	Preferred allocation		Kemeny distance	
	All subjects	Rich and median subjects	All subjects	Rich and median subjects
Perceived deliberation quality	0.256	0.669**	0.040	0.118**
	(0.208)	(0.314)	(0.032)	(0.052)
Social open mindedness	0.200	-0.502	0.028	-0.074
	(0.324)	(0.440)	(0.053)	(0.074)
BF - Negative emotions	0.265	0.068	0.019	-0.030
	(0.192)	(0.265)	(0.030)	(0.043)
BF - Extraversion	0.208	0.141	0.015	0.029
	(0.187)	(0.260)	(0.029)	(0.042)
BF - Open mindedness	-0.042	-0.116	0.021	0.002
	(0.218)	(0.274)	(0.035)	(0.047)
BF - Agreeableness	0.256	-0.287	0.031	-0.073
	(0.233)	(0.317)	(0.037)	(0.054)
BF - Conscientiousness	0.032	0.224	-0.016	0.047
	(0.248)	(0.334)	(0.039)	(0.054)
Machiavelism	0.015	0.035	0.003	0.006
	(0.028)	(0.038)	(0.005)	(0.006)
Female	0.121	0.463	0.003	0.059
	(0.396)	(0.549)	(0.064)	(0.093)
Age	-0.008	-0.059	-0.002	-0.004
	(0.037)	(0.045)	(0.006)	(0.007)
Student	-0.378	-2.165*	0.030	0.042
	(0.791)	(1.275)	(0.128)	(0.179)
Student in Econ & Manag.	-0.420	-0.522	-0.001	-0.011
	(0.410)	(0.525)	(0.067)	(0.088)
Poor & bottom middle class	(ref)	(ref)	(ref)	(ref)
Middle class	1.057**	1.406**	0.089	0.112
	(0.519)	(0.713)	(0.081)	(0.115)
Rich & upper middle class	0.279	0.558	-0.012	0.012
	(0.525)	(0.713)	(0.081)	(0.119)
Constant			-0.520	-0.486
			(0.446)	(0.591)
Observations	160	96	160	96
R^2 or pseudo-R^2	0.077	0.156	0.058	0.153

Notes: *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level, based on an ordered logit model for the preferred allocation in columns 1 and 2 and an OLS for the K-distance in columns 3 and 4. The dependent variable in column 1 is equal to -1 if the subject prefers a less egalitarian allocation in period 2; 0 if she reports the same preferred allocation in both periods; and 1 if she prefers a more egalitarian allocation in period 2. The dependent variable in column 2 is the change in the K-distance between periods 1 and 2.

Table A15: Subjects' characteristics and degree of egalitarianism changes between periods 1 and 2

I. Results for the Borda vote

In this appendix, we present the main results when subjects ranked the allocations in the voting position. Figure A2 reports the distribution of the preferred allocation using the ranking given by the subjects in the voting ranking task and Table A16 reports the mean changes in the Kemeny distance with this task. In a nutshell, these results are very similar to the results obtained for the dictator task.

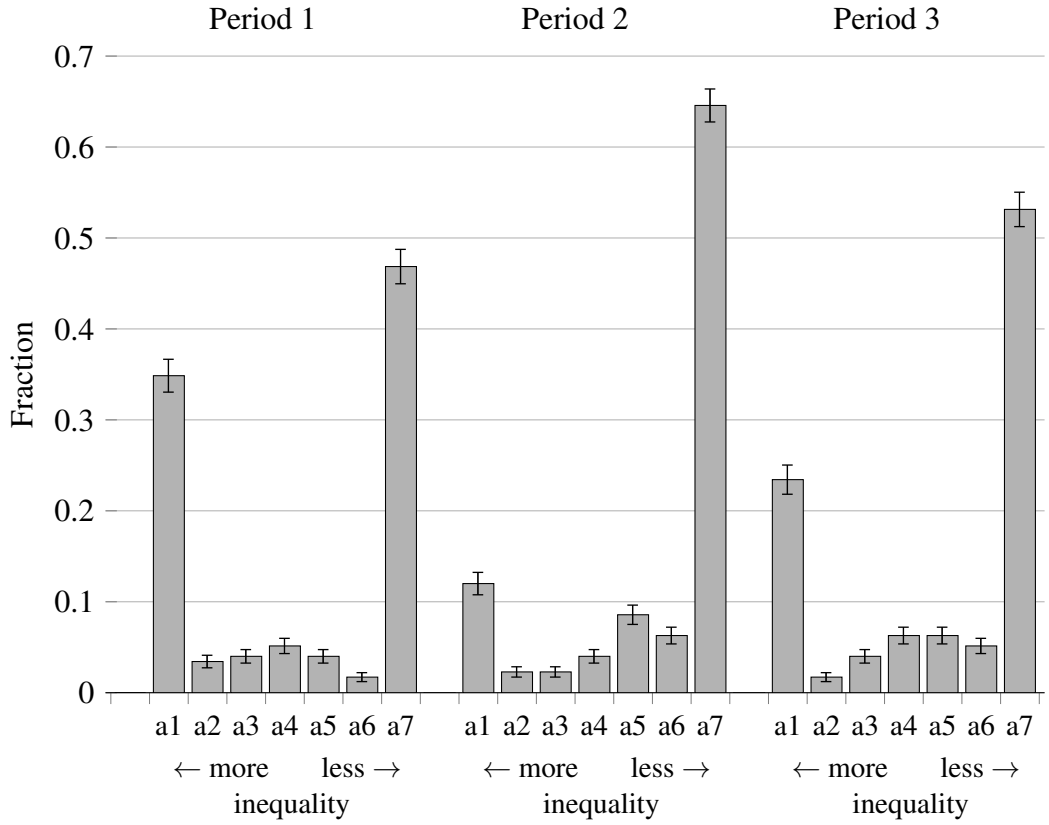

Figure A2: Distribution of preferred allocations in T1 (Borda vote)

Notes: The main bars display percentages of subjects with preferred allocations a1 to a7 per treatment. The error bars display one standard error of the mean, with values based on logit estimates with dummies for allocations for each period.

	Mean	% with K-distance			Signed-rank tests	
		$= 0$	$\in]0,1[$	$= 1$	$=$ period 2	$=$ period 3
Period 1	0.55	30.29	29.71	40.00	$p < 0.001$	$p = 0.022$
Period 2	0.74	9.71	39.43	50.86		$p = 0.031$
Period 3	0.65	19.43	33.71	46.86		

Notes: Signed-rank test for clustered data, clustering at the group level for p1=p2 and at the session level for p2=p3 and p1=p3.

Table A16: Kemeny distances in T1 (Borda vote)

	Total	Identity	Social Comp.	Persuasion	SC+P
All members					
Pref. allocation (coef)	1.140	0.582	0.108	0.449	0.558
Wald tests (p-values)	0.000	0.000	0.632	0.120	0.000
K-distance (coef)	0.196	0.084	0.087	0.024	0.111
Wald tests (p-values)	0.000	0.000	0.059	0.687	0.000
Rich members					
Pref. allocation (coef)	2.426	1.360	0.121	0.945	1.066
Wald tests (p-values)	0.000	0.000	0.747	0.065	0.002
K-distance (coef)	0.299	0.180	0.094	0.025	0.119
Wald tests (p-values)	0.000	0.000	0.149	0.766	0.006
Median members					
Pref. allocation (coef)	1.551	0.204	0.426	0.922	1.348
Wald tests (p-values)	0.000	0.418	0.469	0.147	0.000
K-distance (coef)	0.258	0.014	0.158	0.086	0.245
Wald tests (p-values)	0.000	0.554	0.086	0.387	0.000
Poor members					
Pref. allocation (coef)	0.484	0.343	0.921	-0.780	0.141
Wald tests (p-values)	0.215	0.229	0.190	0.305	0.686
K-distance (coef)	0.061	0.024	0.044	-0.007	0.037
Wald tests (p-values)	0.035	0.250	0.038	0.839	0.225

Notes: This table reports the decomposition results using the regression estimates reported in Table A18 (columns 2 and 4) and Table A19. SC+P stands for the joint effect of social comparison and persuasion.

Table A17: Decomposition results (Borda vote)

Table A17 reports the results of the decomposition of the group deliberation effect with the Borda vote. Results are again very similar to the ones with the dictator task. Social identity has a statistically significant effect on subjects' degree of egalitarianism and explains about half of the effect of group deliberation. As with the dictator task, it is very important for rich members and not statistically significant for median and poor members. Finally, the joint impact of social comparison and persuasion is again highly statistically significant for all subjects, rich and median members.

	(1)	(2)	(3)	(4)
	Preferred allocation		Kemeny distance	
	Unconstrained	Constrained	Unconstrained	Constrained
T1 × Period 1	(ref)	(ref)	(ref)	(ref)
T1 × Period 2	1.057***	1.140***	0.186***	0.196***
	(0.157)	(0.155)	(0.027)	(0.028)
T1 × Period 3	0.471***	0.558***	0.101***	0.111***
	(0.129)	(0.138)	(0.024)	(0.027)
T2 × Period 1	-0.762	0	-0.071	0
	(0.553)		(0.109)	
T2 × Period 2	-0.482	0.108	0.037	0.087*
	(0.545)	(0.226)	(0.107)	(0.046)
T2 × Period 3	-0.314	0.108	0.049	0.087*
	(0.544)	(0.226)	(0.106)	(0.046)
Observations	720	720	720	720
R^2			0.074	0.073
Log Likelihood	-918.382	-920.194		

Notes: This table reports the regression estimates of eq. 5 without any constraints (columns 1 and 3) and assuming $\beta_{12} = 0$ and $\beta_{22} = \beta_{23}$ (columns 2 and 4). *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level, based on an ordered logit model for the preferred allocation in columns 1 and 2 and an OLS for the K-distance in columns 3 and 4, with clustered standard errors per subject and session dummies.

Table A18: Regression estimates (Borda vote)

	(1)	(2)	(3)	(4)	(5)	(6)
	Rich members		Median members		Poor members	
	Pref. alloc.	K-distance	Pref. alloc.	K-distance	Pref. alloc.	K-distance
T1× Period 1	(ref)	(ref)	(ref)	(ref)	(ref)	(ref)
T1× Period 2	2.426***	0.299***	1.551***	0.258***	0.484	0.061**
	(0.354)	(0.048)	(0.430)	(0.060)	(0.390)	(0.029)
T1× Period 3	1.066***	0.119***	1.348***	0.245***	0.141	0.037
	(0.337)	(0.043)	(0.363)	(0.057)	(0.348)	(0.030)
T2 × Period 2 / 3	0.121	0.095	0.426	0.158*	0.921	0.044**
	(0.374)	(0.065)	(0.588)	(0.092)	(0.702)	(0.021)
Observations	288	288	144	144	288	288
R^2		0.230		0.244		0.109
Log Likelihood	-373.952		-147.145		-213.973	

Notes: This table reports the regression estimates of eq. 5 assuming $\beta_{12} = 0$ and $\beta_{22} = \beta_{23}$ and per rank groups. *Significant at 10% level, **Significant at 5% level, ***Significant at 1% level, based on an ordered logit model for the preferred allocation in columns 1, 3, and 5 and an OLS for the K-distance in columns 2, 4 and 6, with clustered standard errors per subject and session dummies.

Table A19: Regression estimates per rank group (Borda vote)

J. Instructions

In this appendix, we present the brief instructions that subjects read at the beginning of the experiment.[17]

Page 1

Welcome

You are all going to participate today in an experiment on decision-making.

This experiment has been designed in such a way that your anonymity is assured.

You will already receive an initial sum of 5 euros for showing up on time and a sum of 2 euros for having answered the online questionnaire. During the experiment, you will be able to obtain additional gains. These gains will depend on your decisions as well as those of the other participants.

The following instructions give you the information you need to participate in the experiment.

Thank you for your participation!

Page 2

The general course of the experiment

The experiment is divided into **3 periods**. During each of these periods, you will be assigned to a group of 5 participants: you and 4 other participants, who will not always be the same. These groups will be formed randomly. In each period, you will be asked to make choices between different **distributions of monetary gains** within your group. These distributions are distinguished according to the total gain and its distribution (more or less equal) between the members of the group.

Your position (more or less favored) within each group is randomly determined.

Your decisions are private and the periods are independent of each other: your decisions in one period do not affect your decisions in the other periods.

At the end of the experiment, we will ask you to answer a few questions.

How is the distribution of gains chosen?

[17] The instructions were given in paper and are translated from French. French instructions are available upon request.

At each period, there will be a choice between **7 possible distributions** of gains.

These will be communicated to you during the experiment and you will be able to see them on your screen when you have choices to make.

There are two possible procedures for deciding the distribution of gains within your group:

1 **Either the distribution is decided by one of the members of the group, and only by him/her.**

2 **Either the distribution is decided by a vote.**

The selection procedures are presented in detail below.

Page 3

The distribution is decided by one of the members of the group, and only by him/her.

In this case, each member of the group has a 1 in 5 chance of being the member who decides which distribution of gains is chosen.

We will therefore ask you to **rank the 7 possible distributions if you are drawn to be the sole decision maker of the distribution of gains within your group.**

If you are drawn to be the sole decision maker of the distribution, the distribution is chosen as follows:

Rank you assign to the distribution	**Chances of the distribution being chosen**
1st	38 out of 100 chances
2nd	24 out of 100 chances
3rd	17 out of 100 chances
4th	11 out of 100 chances
5th	7 out of 100 chances
6th	3 out of 100 chances
7th	0 out of 100 chances

Page 4

The distribution is decided by a vote.

In this case, the 5 members assign points to each distribution and the distribution with the most points is chosen.

We will therefore ask you to **rank the 7 distributions if the vote decides the distribution of gains within your group.**

The number of points you assign to each of the distributions depends on the ranking which you give it as follows:

Rank you assign to the distribution	**Number of points you assign to the distribution**
1st	6 points
2nd	5 points
3rd	4 points
4th	3 points
5th	2 points
6th	1 point
7th	0 points

If several distributions end with the same total number of points, then the chosen distribution will be the one that has been ranked the most often above the others.

Page 5

The final remuneration

Your final remuneration as well as that of the other participants is determined as follows:

1 **One of the 3 periods is chosen randomly.** Each period has a 1 in 3 chance of being selected.

2 For the selected period:

   a There is a **50/50 chance** that the distribution will be decided by one of the members of the group, and only by him/her. ⇒ In this case, a member is randomly selected. **Each member has a 1 in 5 chance of being selected.** It is his/her choices that determine the final remuneration of all members of the group.

   b There is a **50/50 chance** that the distribution will be decided by vote. ⇒ In this case, it is the result of the vote that determines the final remuneration of all the members of the group.

We will let you know your final compensation once the experiment is over. You will also be able to know the selection procedure selected. If the final remuneration is decided by one

of the members of the group, and only by him/her, the identity of this participant will not be revealed.

Before starting the experiment, we will ask you to answer a few comprehension questions.

K. Pre- and post-experimental questionnaires

The perceived quality of deliberation

To measure the perceived quality of the group deliberation, we use a short post-experimental questionnaire where subjects are asked to evaluate the following statements (translated from French):

1. All participants had an equal opportunity to discuss.
2. Participants were respectful of each other.
3. Participants tried to justify their points of view.
4. Participants considered the opinions of others.
5. Participants reconsidered their own beliefs and opinions.
6. Participants were sincere.

The exact wording of the question they are asked to answer is: *Please indicate the extent to which you agree with these statements*. They have to give a score between 1 (*Not at all in agreement*) and 7 (*Totally agree*).

The Big-Five

The remaining questionnaires were completed online before the experiment. First, we use the 15-item version of the Big-Five questionnaire (Gerlitz and Schupp 2005). Subjects are asked to evaluate the following statements:
I see myself as someone who...

1. Works conscientiously (Conscientiousness)
2. Is talkative (Extraversion)

3. Is sometimes rude to others (Agreeableness)
4. Creative, full of original ideas (Open mindedness)
5. Worries a lot (Negative emotions)
6. Is indulgent by nature (Agreeableness)
7. Tends to be lazy (Conscientiousness)
8. Is sociable, outgoing (Extraversion)
9. Enjoys artistic and aesthetic activities (Open mindedness)
10. Is easily anxious (Negative emotions)
11. Is efficient at work (Conscientiousness)
12. Is reserved (Extraversion)
13. Is considerate and kind to almost everyone (Agreeableness)
14. Has a great imagination (Open mindedness)
15. Remains calm in stressful situations (Negative emotions)

The exact question is: *For each statement, indicate the extent to which you agree or disagree with the statement*. For each statement, subjects have the choice between five responses: *Disagree strongly*; *Disagree a little*; *Neutral – no opinion*; *Agree a little*; *Agree strongly*. For the calculation of the mean values, note that we recode answers given to some items (questions 3, 7, 12 and 15).

Machiavelism

Second, subjects are asked to evaluate the following statements taken from the MACH-IV test of Machiavellianism (see, e.g., Christie and Geis 1970):

1. Never tell anyone the real reason for your actions, unless it will help you
2. The best way to behave is to tell people what they want to hear.

3. You should only act when you are sure you are acting morally.
4. Most people have a kind and generous heart.
5. It is best to assume that everyone has a core of badness just waiting to show up.
6. In any situation, it is best to behave with honesty.
7. Lying is never justified.
8. In general, we only work hard if we have to.
9. Humility and honesty are preferable to arrogance and dishonesty.
10. If you ask someone for a favor, it is better to be honest about your reasons than to present other reasons that would carry more weight.
11. For the most part, those who succeed in life are honest and moral.
12. Anyone who relies entirely on others is tempting the devil.
13. The main difference between most criminals and others is that they are stupid enough to get caught.
14. Most people are brave.
15. It is prudent to flatter those in power.
16. Perfect goodness is of this world.
17. It is not true to say that the world is populated by simpletons.
18. It is not easy to go through life without bending the rules.
19. Those who suffer from an incurable disease should be allowed to end their lives painlessly.
20. Most people forget the death of their father more easily than the loss of their possessions.

For each statement, subjects have the choice between five responses: *Disagree strongly*; *Disagree a little*; *Nor disagree, nor agree*; *Agree a little*; *Agree strongly*.

Social open mindedness

Finally, subjects are asked to evaluate the following statements related to (social) open mindedness:

1. There are two kinds of people in the world: those who are for the truth and those who are against it.

2. It is dangerous to compromise with our political opponents, because it usually leads to betrayal of our own side.

3. A group that tolerates too many differences of opinion among its own members cannot exist for long.

The question subjects are asked to answer: *To what extent do you agree or disagree with these statements?* Subjects have the choice between five responses: *Disagree strongly*; *Disagree a little*; *Nor disagree, nor agree*; *Agree a little*; *Agree strongly*.