Supplementary material

A Theoretical appendix

A.1 The game

There are two states of the world, X and Y, and six players that form a group G, with three white and three blue players forming two respective subgroups, G_w and G_b. A player's color is publicly observable. The players have to choose a policy P from three alternative policies, A, B, and C, by a vote. Policies generate state-dependent payoffs that may differ across colors. These payoffs are depicted in Table 1 in the main text. Nature draws the state of the world ω, which is either X or Y with equal probability, at the beginning of the game. Afterwards, nature randomly draws an informative private signal $s_i \in \{x, y\}$ on the state of the world for each white player i and sends an empty signal $s_i = \emptyset$ to the blue players. Informative signals are conditionally independent and true with probability $p := \{s_i = x \mid \omega = X\} = \{s_i = y \mid \omega = Y\} = 0.7$. The subsequent collective policy choice has two stages, the communication stage and the voting stage, in treatments *FullyPublic*, *TopDown* and *TopDownClosed*. Treatment *NoChat* has no communication stage.

The voting stage is identical across treatments and is structured as follows: All six players simultaneously and individually place a vote for A, B, or C, or abstain. The winning alternative is determined by the plurality rule, i.e., the alternative with the most votes is implemented. If there is a tie, the winning alternative is chosen randomly, with equal probability of both alternatives. In the end, payoffs from the winning alternative are realized, given the true state of the word.

In all treatments with communication stage, this stage is structured as follows: There is a set of senders $S \supseteq G_\tau$ and a set of receivers $R(G_\tau)$ to which players in the subgroup G_τ, $\tau \in \{w, b\}$, can send messages. Let M denote the set of all messages that can be constructed in the common language spoken by the six players, including the empty set. Then, any player $i \in G_\tau \subseteq S$ sends a message $m_i \in M$ to $R(G_\tau)$. In *FullyPublic*, $S = G = R(G_w) = R(G_b)$, i.e., communication is public and involves everyone as both sender and receiver. In particular, the whites may reveal their signal to the entire group of six (or lie or be silent about it), and both the whites and the blues may recommend a specific voting profile for the group. *TopDown* differs from *FullyPublic* in that $S = G_w$, $R(G_w) = G$, and $R(G_b) = \emptyset$. Thus, blues are no longer senders, but messages are still received by everyone. In *TopDownClosed*, by contrast, $S = G_w = R$ on the first communication stage, i.e., the blues are entirely excluded from the communication on that stage, and the whites send messages to the subgroup of white players only. On the second communication stage in *TopDownClosed*, the whites can talk to the entire group; hence, $R = G$.

A.2 Preferences

According to our main underlying assumption, a player's expected utility is the expected sum of his own payoff and the payoffs of his receivers on the – first – communication stage. Hence, players have treatment-dependent preferences that can be described as follows. Let $R^1(G_\tau)$ denote the set of receivers of individuals in G_τ on the first communication stage in the game (which is also the final communication stage in all treatments except *TopDownClosed*). Let $\pi_i(\omega, P)$ denote the final payoff of player i, given the state of the world ω and the chosen policy P. Moreover, let $q_i(\omega \mid m, s_i)$ denote player i's posterior belief about how likely state of the world ω is, given the sent messages m and his own signal s_i; let $\sigma_P(v) \in \left\{0, \frac{1}{2}, 1\right\}$ denote the probability of P being the winning policy, given the voting profile v; and let $z(P) \in \{1, 2, 3\}$ be an index of the policy and $z(\omega) \in \{1, 2\}$ an index of the state of the world. Hence, we can define the utility function as follows:

$$u_i(v \mid m, s_i) = \sum_{z(\omega)=1}^{2} q_i(\omega \mid m, s_i) \sum_{z(P)=1}^{3} \sigma_{P_{z(P)}}(v) \left(\pi_i\left(\omega, P_{z(P)}\right) + \sum_{\substack{j \in R^1(G_\tau) \\ i \in G_\tau}} \pi_j\left(\omega, P_{z(P)}\right) \right).$$

A.3 Equilibrium concept

We are solving the game for all Perfect Bayesian Nash equilibria in pure strategies that fulfill the following selection criteria.

Definition 1 *(**WU**) Any equilibrium is in **weakly undominated** strategies.*

Definition 2 *(**DT**) Players exhibit **dominant truth-telling**: If there exists a truth-telling equilibrium, no babbling equilibrium is played; i.e., if there exists an equilibrium in which all whites reveal their signal on the / a communication stage, no equilibrium is played in which not all whites reveal their signal on that stage.*

Definition 3 *(**SCT**) Players exhibit **same-color trust**: If the message of a player i to the entire group contradicts a message he has sent to players of his own color only, players of the same color as i believe the message that i has sent to them and disbelieve the message he has sent to the entire group.*

Definition 4 *(**MC**) Players exhibit **minimal coordination** in the following sense: For any $\tau \in \{w, b\}$, players $i \in G_\tau$ who move at the same information set I and hence know* *that they have identical beliefs $q_i(\omega \mid I) = q$ coordinate on sending the same message and / or voting for the same policy such that they maximize their (joint and individual) expected utility, given the strategies of the other voters.*

Definition 5 *(**LS**) Whites exhibit **literal speaking**: They send the message "x" if they want to indicate that their signal was "x", and they send the message "y" if they want to indicate that their signal was "y".*

Definition 6 *(**CORS**) All players exhibit **conditioning on revealed signals**: They may condition their strategies on the signals that are revealed by the messages but do not use messages as coordination devices otherwise.*

Criterion **WU** excludes equilibria in which *selfish* whites vote for policy C. Criterion **DT** is typical for the cheap-talk literature and selects the equilibrium with the highest degree of information transmission. Criterion **SCT** exludes equilibria in *TopDownClosed* in which the whites cannot lie to the blues without changing the beliefs of the other whites, too. (Note that such equilibria would be extremely implausible since in *TopDownClosed*, the whites can even tell each other that they intend to lie to the blues.) Criterion **MC** restricts attention to equilibria in which the blues coordinate on the same voting strategy (since the blues always have the same information). Moreover, **MC** guarantees that in any truth-telling equilibrium (in which the whites, too, have the same information) the whites also coordinate on the same voting strategy. Criterion **LS** reduces the syntax of the language in which signals are communicated to a binary set and hence simplifies (the proofs in) equilibrium description. **CORS** restricts the function of communication as a coordination device to what is implied by **MC** and allows us to focus on information aggregation rather than pure (uninformed) coordination. The resulting effect of **CORS** is to restrict the number of outcome equivalent equilibria that differ in strategy profiles.

We now present the equilibria of the different treatments in the order of decreasing inclusiveness (FullyPublic, TopDown, TopDownClosed, NoChat).

A.4 Equilibria in *FullyPublic*

Since in *FullyPublic* both the whites and the blues send messages to the entire group and care for both colors, i.e., they have efficiency preferences. Hence, their (joint) utility is

$$
\begin{aligned}
u_i^P(v) &= \sum_{z(\omega)=1}^{2} q_i(\omega \mid m, s_i) \sum_{z(P)=1}^{3} \sigma_{P_{z(P)}}(v \mid \omega) \left(\pi_i\left(\omega, P_{z(P)}\right) + \sum_{j \in G} \pi_j\left(\omega, P_{z(P)}\right) \right) \\
&= q_i(X \mid m, s_i)\left(\sigma_A(v \mid X) 120 + \sigma_C(v \mid X) 30\right) + \\
&+ q_i(Y \mid m, s_i)\left(\sigma_A(v \mid Y) 30 + \sigma_B(v \mid Y) 90 + \sigma_C(v \mid Y) 60\right).
\end{aligned}
$$

Consider a candidate equilibrium in which the whites truthfully reveal their signals to the public. In such an equilibrium, all players have the same information and hence then same belief about ω: $q_i(\omega \mid m, s_i) = q(\omega \mid m) \forall i, \omega$. Hence, MC applies to both whites and blues: Whites coordinate on the same vote and blues coordinate on the same vote, maximizing the expected group payoff, given the strategy of the other color. Strategies of the whites may condition on the private signal or on the messages. Note that under truth-telling, the state of the world that is more likely than the other is indicated both by the signal that is received most often (i.e., twice or even three times) and by the message that is sent most often. Hereafter, we will call this signal and message the *majority signal* and the *majority message.*

Definition 7 *An equilibrium is **efficient** if and only if A is the winning policy whenever the majority signal is X and B is the winning policy otherwise.*

Proposition 1 *(i) There is a set of truth-telling equilibria in FullyPublic that fulfill the selection criteria. They have the following properties: (a) The whites reveal their signals. The whites vote for A if the majority message indicates X and for B otherwise, and the blues abstain (LTED: "let the experts decide"). If (off equilibrium) there is no majority message, then arbitrary off-equilibrium beliefs about the unrevealed signal and voting profiles consistent with these beliefs can be assumed. (b) The whites reveal their signals. Everyone votes for A if the majority message indicates X and votes for B otherwise (AA/BB). If (off equilibrium) there is no majority message, then arbitrary off-equilibrium beliefs about the unrevealed signal and voting profiles consistent with these beliefs can be assumed. (ii) These equilibria are outcome equivalent in the sense that A is the winning policy if the majority signal indicates X, and B is the winning policy otherwise. (iii) Hence, both types of truth-telling equilibria are efficient. (iv) These equilibria are the unique pure-strategy equilibria in FullyPublic that fulfill all selection criteria.*

Proof. We first show that the voting profiles described in (a) and (b) are efficient equilibria of the continuation game on the voting stage, given truth-telling. We then show that truth-telling is an equilibrium strategy of the whites, given the voting profiles in (a) and (b). Finally, we prove (i), i.e., that the two equilibria are the only truth-telling equilibria that fulfill our selection criteria. Consider (a) and (b). Due to truth-telling, the majority message always equals the majority signal. Hence, *LTED* and *AA/BB* are efficient. Efficiency preferences imply that no-one has a deviation incentive. Thus, the voting profiles in (a) and (b) are equilibria of the continuation game on the voting stage. Consider now the communication stage preceding the voting stage with *LTED.* If a white deviates from telling the truth in that he lies about his signal, then he either does not change the majority message, hence leaving the voting outcome unchanged, too, or he changes the majority message and hence moves the voting outcome away from efficiency. Thus, no white wants to deviate to lying. Now consider a deviation to silence. Again, this either changes nothing or moves the voting outcome away from efficiency, depending on the off-equilibrium voting strategies. Hence, again, the whites do not want to deviate. The same kind of argument holds true for the communication stage that precedes a voting profile described in (b). Thus, in *FullyPublic* there exist truth-telling equilibria with voting profiles as described in (a) and (b). Parts (ii) and (iii) follow directly.

It now remains to show that these two sets of equilibria defined above contain the only truth-telling equilibria in *FullyPublic* that fulfill our selection criteria. Note first that *CORS* excludes equilibria in which strategies condition on messages without conditioning on beliefs. Furthermore, note that under truth-telling, *MC* applies both to the whites and the blues. If the whites have revealed their signals and the blues abstain, then *MC* and efficiency preferences imply that the whites coordinate on voting for A or B, depending on the majority message. If the whites do this, and if they have revealed their signals, then *MC* and efficiency preferences imply that the blues coordinate on a strategy that never distorts the voting outcome away from efficiency. Hence, in this case the only two voting profiles of the blues that fulfill *MC* are abstention and voting along with the whites. Finally, note that *DT* excludes equilibria with partial truth-telling. Hence, *CORS*, *MC,* *DT,* and efficiency preferences pin down all truth-telling equilibria in *FullyPublic* to the ones that are characterized in (a) and (b). Part (iv) follows directly from this and *DT*. ∎

From Proposition 1, the following result can be derived:

Result 1 *In FullyPublic, (a) the whites truthfully reveal their signal; and (b) if the majority signal indicates X, all votes that are placed are for A; whereas (c) if the majority signal indicates Y, all votes that are placed are for B (AA/BB).*

A.5 Equilibria in *TopDown*

In *TopDown*, the whites can still address the entire group on the communication stage, but the blues are no longer senders. Hence, the whites still have efficiency preferences, but the blues become self-interested. Still, the blues have the same information, so MC still applies to them: $q_i(\omega \mid m) = q(\omega \mid m)\ \forall i \in G_b$. Moreover, note that payoffs are perfectly aligned across players of the same color; thus we can define $\pi_i\left(\omega, P_{z(P)}\right) := \pi_b\left(\omega, P_{z(P)}\right)$ $\forall\ i \in G_b$. Hence, in *TopDown* a player i has utility u_i^{TD} as follows:

$$u_i^{TD} = u_i^P(v) \ \text{ if } i \in G_w,$$

$$u_i^{TD} = \sum_{z(\omega)=1}^{2} q(\omega \mid m) \sum_{z(P)=1}^{3} \sigma_{P_{z(P)}}(v \mid \omega)\, \pi_b\left(\omega, P_{z(P)}\right) \ \text{ if } i \in G_b.$$

Consider truth-telling equilibria.

Proposition 2 *(i) There is a set of truth-telling equilibria in TopDown that fulfill the selection criteria. They have the following properties: The whites reveal their signals. If the majority message indicates X, then (a) all vote for A, (b) all whites vote for A and all blues abstain, or (c) all whites abstain and all blues vote for A. If the majority message indicates Y, then the whites vote for B and the blues for C. If (off equilibrium) there is no majority message, we restrict off-equilibrium beliefs as follows: If there is a one-shot deviation of one white player to being silent and the remaining revealed signals contradict each other(i.e., there is no majority message), then the blues have a belief $q(X \mid m) < \frac{2}{3}$. Then in all voting profiles consistent with off-equilibrium beliefs after such a deviation, the blues vote for C. (ii) These equilibria are outcome equivalent: They generate winning policy A if the majority signal is X and a tie between B and C if the majority signal is Y. (iii) These equilibria are inefficient. (iv) These equilibria are the unique pure-strategy equilibria that fulfill all selection criteria.*

Proof. We first show that with truth-telling of the whites on the communication stage, voting profiles with the properties described in (i) are equilibria of the continuation game. Second, we show that then, truth-telling must be part of the equilibrium. Part (ii) directly follows from part (i); and (iii) directly follows from (ii) and the definition of efficient equilibrium.

Assume now truth-telling of the whites, and consider the blues first. For $q(\omega \mid m) < \frac{2}{3}$, policy C is strictly better for a blue player than the other policies, otherwise, policy A is better than the other policies. If x is the majority message, we have $q(\omega \mid m) \geq 0.7 > \frac{2}{3}$, and if y is the majority message, we have $q(\omega \mid m) \leq 0.3 < \frac{1}{3}$. Thus, if x is the majority message, then A is better for any blue player than (a tie with) any other policy; and if y is the majority message, then (a tie with) C is better for any blue player than (a tie with) any other policy. Then, MC implies that all blues vote for A or abstain if x is the majority message and vote for C otherwise. If there is no majority message (i.e., there are only two messages that contradict each other), then the off-equilibrium belief of the blues, $q(\omega \mid m) < \frac{2}{3}$, and MC imply that all blues vote for C.

Consider now the whites on the voting stage. Remember that they have efficiency preferences. If the majority message is x, then any white prefers A over all other policies. MC then implies that all whites coordinate on an action that makes A the winning policy; i.e., voting for A, or, (only) if the blues vote for A, abstention. If the majority message is y, then any white anticipates the three blue votes for C but prefers B over all other policies himself. Hence, he also prefers a tie between B and C over C or any other tie with C. Thus, MC implies that all whites vote for B. If there is no majority message, i.e., if there are only two messages that contradict each other, then any off-equilibrium belief about the unrevealed signal and any consistent voting strategy of the whites can be assumed. Note that regardless of the voting strategy of the whites after such a deviation, the resulting efficiency level (group payoff) cannot exceed the level implied by the equilibrium strategies (because strategies cannot improve upon conditioning on the full information about all signals). Thus far, we have shown that under truth-telling, voting profiles with the properties described in (a), (b), and (c) are equilibria of the continuation game on the voting stage.

Consider now the communication stage. We check the incentive of an arbitrary white player i to deviate to a lie or to being silent about his signal. Consider now a white who has received a signal s_i. If he lies or is silent about s_i, then he is either not pivotal, the other two messages being $m_{-i} = (y, y)$ or $m_{-i} = (x, x)$, in which case the deviation does not change anything. Or i is pivotal, in which case the other two whites have contradicting signals and s_i is the majority signal. Then, i's efficiency preferences imply that he cannot do better than revealing his signal. Thus, there is no deviation incentive on the communication stage.

We now proceed to proving (i) by showing that *all* truth-telling equilibria in *TopDown* have the properties that are described in (a), (b), and (c). Note first that *CORS* excludes equilibria in which strategies condition on messages without conditioning on beliefs. Second, under truth-telling, MC applies to both colors. Hence, under truth-telling each color will coordinate on an action that maximizes the probability of the policy preferred by this color, given the strategy of the other color and the common beliefs about the state of the world. But then, (a), (b), and (c) describe all voting profiles under truth-telling. Moreover, DT excludes partial truth-telling and babbling equilibria. Thus, MC, *CORS,* and

DT restrict all pure-strategy equilibria in *TopDown* to the set described in Proposition 2, which proves part (iv). ■

From Proposition 2, the following result can be derived:

Result 2 *In TopDown, (a) the whites truthfully reveal their signal; and (b) if the majority signal indicates X, all votes that are placed are for A whereas (c) if the majority signal indicates Y, the whites vote for B and the blues for C (AA/BC).*

A.6 Equilibria in *TopDownClosed*

In *TopDownClosed*, our main underlying assumption implies that the whites do not have efficiency preferences any longer but maximize the joint payoffs of their own color group instead (color preference type). Note that this is equivalent to being selfish since payoffs are perfectly aligned between individuals of the same color. Importantly, WU implies that selfish whites never vote for C, since they prefer any possible outcome of the vote over C, regardless of their beliefs about the state of the world, so that voting for C is a weakly dominated strategy for selfish whites. The blues, too, are selfish, as in *TopDown*.

Consider now potential equilibria in which the whites truthfully reveal their signals to each other on the first communication stage. Note that in such equilibria, the whites have identical beliefs on the voting stage, so that MC applies to them. Note that MC always applies to the blues, regardless of whether they are told the true signals or not.

Proposition 3 *(i) There is a set of equilibria in TopDownClosed that fulfill the selection criteria. They have the following properties: The whites reveal their signals to each other, but babble to the blues. The whites vote for A if the majority message indicates X and for B otherwise, and the blues vote for C (AC/BC). If (off equilibrium) there is no majority message on the first communication stage, arbitrary off-equilibrium beliefs of the whites and white votes consistent with these beliefs can be assumed; but the blues (unobservant of the deviation) are restricted to keep their prior beliefs and hence to vote for C. (ii) These equilibria are inefficient. (iii) These equilibria are the unique pure-strategy equilibria that fulfill all selection criteria.*

Proof. We first show that given truth-telling on the first communication stage, there can be no truth-telling on the second communication stage. We then show existence of the AC/BC equilibria as characterized in (i). Part (ii) – inefficiency – directly follows from (i) and the definition of efficiency. Finally, we will prove (iii).

Assume now that the whites truthfully reveal their signals to each other on the communication stage. Assume for the sake of argument that there is also truth-telling on the second communication stage. Consider now a situation in which the majority signal indicates Y, but there has been one signal indicating X. On the voting stage, both the whites and the blues hence believe that the state of the world is Y with probability 0.7. But then, their preferences and *MC* imply that the whites vote for B and the blues for C. Under truth-telling, the whites' expected utility is $0.3 \times 0 + 0.7 \left(\frac{1}{2} \times 20 + \frac{1}{2} \times 0\right) = 7$. If, by contrast, one of the whites who have received the signal indicating Y deviates to a lie, saying that his signal indicates X, the beliefs of the whites will not change since this is precluded by *SCT*, but the blues will believe that the state of the world is X with probability 0.7. Then, *MC* and the players' preferences imply that the whites will still vote for B and the blues will vote for A. For the whites, this yields an expected utility of

$$0.3\left(\frac{1}{2} \times 20 + \frac{1}{2} \times 0\right) + 0.7\left(\frac{1}{2} \times 10 + \frac{1}{2} \times 20\right) = 13.5.$$

Thus, the lie strictly increases the expected utility of the whites. Consider now a white i whose signal was $s_i = Y$. This white is pivotal on the second communication stage in the sense that his message determines the majority message sent to the blues (since the other two whites are assumed to tell their true – contradictory – signals). Thus, this white has a strict incentive to lie on the second communication stage. This proves that under truth-telling on the first communication stage, there can be no truth-telling on the second communication stage if the signal distribution is $2:1$.

Consider now a situation in which all three signals indicate Y. Then, under truth-telling on both communication stages, no white is the pivotal sender on the second communication stage any longer, and the individual lying incentive does no longer exist on the equilibrium path. Instead, a given white in this situation is indifferent between lying and revealing his signal, given that the other two whites reveal that their signals indicated Y. (Note that the whites know the signal distribution on the second communication stage since we assume truth-telling on the first communication stage.) However, the whites still *prefer* that the blues vote for A rather than C. To see this, note that their expected utility if the blues vote for A (and they themselves for B) would be

$$\frac{0.3^3}{0.3^3+0.7^3}\left(\frac{1}{2} \times 3 \times 20 + \frac{1}{2} \times 3 \times 0\right) + \frac{0.7^3}{0.3^3+0.7^3}\left(\frac{1}{2} \times 3 \times 10 + \frac{1}{2} \times 3 \times 20\right)$$
$$= 43.905.$$

By contrast, if the blues vote for C, the whites' expected utility amounts to

$$\frac{0.3^3}{0.3^3+0.7^3}\times 0+\frac{0.7^3}{0.3^3+0.7^3}\left(\frac{1}{2}\times 3\times 20+\frac{1}{2}\times 3\times 0\right)=27.811.$$

Thus, the whites have a higher expected utility if the blues vote for A rather than C. Now note that the whites have identical beliefs on the second communication stage due to truth-telling on the first communication stage. Thus, *MC* applies to them on the second communication stage. But sending a majority message that indicates Y and thus making the blues vote for C violates *MC*. Thus, our selection criteria exclude equilibria in which any signal distribution leads to truth-telling on the second communication stage.

Consider now potential equilibria with truth-telling on the first communication stage and babbling on the second communication stage. Consider the voting stage first. The blues have their prior belief that both states of the world are equally likely. Thus, their selfish preferences and *MC* imply that they coordinate on voting for C. The whites, by contrast, know the actual signal distribution s. They prefer A whenever $q(X \mid s) \geq 0.7$ and B otherwise. Hence, they also prefer a tie between A and C whenever $q(X \mid s) \geq 0.7$ and a tie between B and C otherwise. But then, *MC* implies that they coordinate on voting for A whenever the majority signal indicates X and on voting for B otherwise. This proves the voting profile AC/BC on the equilibrium path.

Consider now the second communication stage. Given that the blues do not condition their beliefs on the messages sent, no white has an incentive to deviate from babbling to conditioning his message on his signal. Note that this also holds true off equilibrium, i.e., after a deviation of a white / some whites on the first communication stage.

Now consider the first communication stage. Given that the whites believe each other, no white has an incentive to deviate to being silent or to lying. To see this, note that such a deviation would either change nothing or would distort the beliefs of the other two whites away from the true signal distribution. This distortion, in its turn, would either change nothing or distort the votes of the other two whites away from the voting profile that maximizes the whites' expected utility, given that the blues vote for C. Note that the blues cannot observe any deviation on the first communication stage. Hence, they cannot respond to such a deviation and will vote for C after it, too.

This proves parts (i) and (iii) of Proposition 3. Part (ii) is trivial. ■

From Proposition 3, the following result can be derived:

Result 3 *In TopDownClosed, (a) the whites truthfully reveal their signals to each other but babble to the blues, and (b) if the majority signal indicates X, the whites vote for A but the blues for C, whereas (c) if the majority signal indicates Y, the whites vote for B and the blues still for C (AC/BC).*

A.7 Equilibria in *NoChat*

In the *NoChat* treatment, there is no possibility to communicate. Therefore, both colors become self-interested and maximize the utility of their own color. Moreover, only the blues (know that they) have the same information set, namely their prior belief that the two states of the world are equally likely. The whites, however, have private independent information on the true state. Hence, MC applies to the blues but not to the whites.

Proposition 4 *(i) There is a set of equilibria in NoChat that fulfill the selection criteria. They have the following properties: The blues vote for C, and (a) the whites vote for A (AC/AC), or (b) the whites vote for B (BC/BC), or the whites vote for A if their signal indicates X and for B otherwise (split-whites). (ii) These equilibria are inefficient.*

Proof. Since MC applies to the blues, we only have equilibria in which the blues coordinate on the same vote. Since the blues are self-interested in *NoChat*, votes other than C are weakly dominated for them. Hence, MC and WU restrict the analysis to equilibria in which the blues vote for C. The whites are self-interested, too. Given that the blues vote for C, each white will minimize the probability of the implementation of C (since C provides strictly lower expected payoffs than any other policy for a self-interested white, regardless of his signal). Voting for C is hence weakly dominated for the whites. Thus, WU exludes equilibria in which some whites, too, vote for C. Consider now an arbitrary white i. If the two other whites vote for the same policy (that is not C), then i's best response is to vote for this policy, too, in order to decrease the probability of C from 1 to 0.5. Hence, *AC/AC* and *BC/BC* are equilibria. If, now, the two other whites vote for the policy indicated by their signal (A if the signal indicates X and B otherwise), the best response of i is to vote in line with his signal, too. To see this, note that this strategy maximizes the probability of hitting the vote of the other two whites if they voted for the same policy, and hence minimizes the probability of C. Thus, *split-white*s is an equilibrium, too. Note that self-interest, MC (for the blues) and WU (for both colors) exclude other possible equilibria. This proves (i). Part (ii) follows from the definition of efficiency. ■

Proposition 4 implies the following result:

Result 4 *In NoChat, the blues vote for C, regardless of the majority signal; and the whites vote for A or B or according to their signal.*

Excluding all equilibrium outcomes with abstention, results 1-4 imply our testable hypotheses 1-4 below. In these hypotheses, we focus on the treatment comparisons, i.e., on the comparative statics, rather than on point predictions.

Hypothesis 1 (Whites' truth-telling) *The whites report signals truthfully to the same extent in FullyPublic and TopDown and lie comparably more often in TopDownClosed. That is, the frequency ordering of instances where the majority message equals the majority signal is: FullyPublic = TopDown > TopDownClosed.*

Hypothesis 2 (Blues' trustfulness) *The blues believe in the whites' majority message x to the same extent in FullyPublic and TopDown and comparably less in TopDownClosed. That is, the frequency ordering of blues' votes for A after majority message x is: FullyPublic = TopDown > TopDownClosed.*

Hypothesis 3 (Whites' voting decisions) *Given majority signal x (y), the whites vote for A (B) at least as often in FullyPublic, TopDown and TopDownClosed as in NoChat. That is, the frequency ordering of whites' votes for A as well as for B is: FullyPublic = TopDown = TopDownClosed ≥ NoChat. Whites never vote for C.*

Hypothesis 4 (Blues' voting decisions) *Given majority signal x, the blues vote for A more often in FullyPublic and TopDown than in NoChat and TopDownClosed. Given majority signal y, the blues vote for B more often in FullyPublic than in NoChat, TopDown and TopDownClosed. That is, the frequency ordering of blues' votes for A (B) is: FullyPublic = TopDown > NoChat = TopDownClosed (FullyPublic > NoChat = TopDown = TopDownClosed). The frequency orderings of blues' votes for C are reversed, compared to the above frequency orderings.*

A.8 Predictions with standard preferences

Standard preferences would imply that all players are selfish maximizers of their own expected payoff. Due to our design, this is equivalent to assuming a color preference type for both colors in all treatments. Hence, the predictions for treatments *NoChat* and *TopDownClosed* would not change if we assumed standard preferences.

By contrast, our predictions for *FullyPublic* and *TopDown* would change: As is easy to show, there would not be any truth-telling equilibria but only babbling equilibria in these two treatments since each white would have an incentive to lie to the blues and report "x" even if her signal indicated "y". We omit the proof, but a crucial point in the proof is that selfish whites prefer AA over BC even under majority signal y. Accordingly, if players were selfish, the predictions for *FullyPublic* and *TopDown* would coincide with those for *NoChat*.

B Additional empirical analyses

In this appendix we present additional analyses based on the equilibrium predictions of our model presented in Appendix A. We first focus on realized voting outcomes conditional on the majority signal, that is, the signal received by the majority of whites. In Table B.1 we present the frequencies of the following voting outcomes that we introduced in Appendix A above: AA, BB, AC, and BC. As a reminder: AA and BB refer to voting outcomes in which all six voting group members either vote for A or B, respectively. AC and BC describe voting outcomes in which the three blues vote for C and the three whites either for A or B, respectively. We expect AA and AC to only occur if the majority signal is x, and BB and BC only if the majority signal is Y. Besides these voting outcomes, we also consider the *split-whites* outcome in which all blues vote for C and all whites follow their individual signal (vote for A if their own signal is x, vote for B if it is y). Grey cells in Table B.1 indicate the predicted equilibrium outcomes as derived in Appendix A, figures printed in bold highlight observed modal voting outcomes. Besides the precise equilibrium outcomes, we also present information about voting outcomes in which at most one of the blue and/or one of the white players deviates from the equilibrium strategy. We call these realizations “almost” realizations.

In Table B.2, we regress the voting outcomes that are part of predicted equilibria on treatment dummies and additionally control for period effects. In all regressions, *NoChat* serves as baseline treatment. Additional results from Wald tests on treatment differences are presented in the bottom part of the table.

As predicted, we find significantly more AA outcomes and significantly fewer AC outcomes after majority signal x in *FullyPublic* and *TopDown* than in *NoChat*. This becomes already visible in the left part of Table B.1 and is corroborated by the linear probability models (1) and (2) in Table B.2. However, these regressions also reveal deviations from the model's predictions: In case of majority signal x, there are significantly more AA outcomes and significantly fewer AC outcomes in *TopDownClosed* than in *NoChat*, where the predictions were no differences. Relatedly, *TopDownClosed* does not lead to significantly fewer AA outcomes (more AC outcomes) than *TopDown*, other than predicted. Moreover, there are significantly more AA outcomes in *FullyPublic* than in *TopDown*, where the prediction, again, was no difference.

After majority signal y, we find significantly more BB outcomes in *FullyPublic* than in any other treatment, as predicted. This can be seen in the right part of Table B.1 and in Model (3) of Table B.2. We also find that BC occurs significantly more often in *TopDown* and *TopDownClosed* than in *NoChat*, again as predicted (see Model (4)). However, our regressions also reveal deviations, namely that BC does not occurs significantly less often in *FullyPublic* than in the other communication treatments (see Model (4)).

Table B.1: Voting outcomes at the group level – Conditional on the received majority signal

	Majority signal: x				Majority signal: y			
	NoC	FP	TD	TDC	NoC	FP	TD	TDC
AA	0	**0.386**	**0.201**	0.143	0	0.059	0	0.021
Almost AA	0.038	0.284	0.358	0.328	0.018	0.059	0.055	0.105
(Almost) AA	0.038	0.670	0.559	0.471	0.018	0.119	0.055	0.126
AC outcome	**0.338**	0.121	0.106	**0.159**	**0.295**	0.103	0.114	0.052
of this: Split-whites	0.184	0.023	0.011	0.037	-	-	-	-
Almost AC outcome	0.513	0.181	0.307	0.339	0.476	0.086	0.095	0.126
(Almost) AC outcome	0.850	0.302	0.413	0.497	0.771	0.189	0.209	0.178
BB	0	0	0	0	0	0.086	0.005	0.005
Almost BB	0	0.005	0	0.005	0	0.124	0.065	0.058
(Almost) BB	0	0.005	0	0.005	0	0.211	0.070	0.063
BC outcome	0	0	0	0	0.012	**0.216**	**0.313**	**0.283**
of this: Split-whites	-	-	-	-	0.006	0.092	0.164	0.152
Almost BC outcome	0.013	0	0.006	0	0.090	0.238	0.328	0.257
(Almost) BC outcome	0.013	0	0.006	0	0.102	0.454	0.642	0.539
Split-whites	0.269	0.023	0.017	0.037	0.030	0.108	0.174	0.168
Other	0.098	0.023	0.022	0.026	0.108	0.027	0.025	0.094
Observations	234	215	179	189	166	185	201	191

In "Almost" outcomes at most one player per color group deviates from the respective outcome. "(Almost)" outcomes comprise both the "Almost" outcomes and the outcomes itself, without deviation. Grey cells indicate predicted equilibrium outcomes, figures printed in bold highlight the observed modal voting outcomes for the respective treatment and signal combination. Treatments are abbreviated as follows: *FullyPublic*: *FP*, *TopDown*: *TD*, *TopDownClosed*: *TDC*, *NoChat*: *NoC*.

Table B.2: Voting outcomes

	Majority signal: x		Majority signal: y	
	(1) AA	(2) AC	(3) BB	(4) BC
FullyPublic (FP)	0.403***	−0.226***	0.080***	0.212***
	(0.000)	(0.004)	(0.001)	(0.000)
TopDown (TD)	0.213***	−0.238***	0.000	0.307***
	(0.000)	(0.002)	(0.934)	(0.000)
TopDownClosed (TDC)	0.147***	−0.181***	0.003	0.274***
	(0.000)	(0.007)	(0.585)	(0.000)
Period	−0.019***	0.010***	−0.005**	0.007**
	(0.000)	(0.001)	(0.026)	(0.033)
Constant	0.194***	0.233***	0.061**	−0.063*
	(0.000)	(0.001)	(0.026)	(0.088)
Wald test results for comparison of treatment coefficients (p values):				
FP vs. TD	0.004	0.821	0.002	0.139
TD vs. TDC	0.201	0.180	0.722	0.433
FP vs. TDC	0.000	0.282	0.002	0.221
R^2	0.230	0.085	0.095	0.086
Number of clusters	20	20	20	20
Observations	817	817	743	743

Linear probability models. Dependent variable: Realization of the respective voting outcome. Robust standard errors are clustered at the session level and p-values are given in parentheses: * p<0.10, ** p<0.05, *** p<0.01. NoChat serves as baseline treatment in all regressions.

Next, we consider realized earnings and efficiency. Table B.3 summarizes the predicted and actual (expected) joint earnings that are realized by the voting group and the color groups, respectively. Regressions of these expected joint earnings on treatment dummies are presented in Table B.4.

Our efficiency results are more (though also not fully) in line with the model's predictions. Comparing total expected earnings – our measure of efficiency – between the communication treatments in Table B.3, we find that they are highest in *FullyPublic* (80.98) and lowest in *TopDownClosed* (78.28), with *TopDown* (79.21) in between, as predicted. However, the Wald test results from Model (1) presented in the bottom part of Table B.4 reveal that none of the deliberation treatment comparisons in (voting group or color group) expected earnings is statistically significant. The only statistical differences are between the deliberation treatments and *NoChat*. Hence, *FullyPublic* is less and *TopDownClosed* more efficient than predicted, compared to *TopDown*.

Table B.3: Rankings over earnings

Efficiency		White players' joint earnings		Blue players' joint earnings	
Predicted expected total payoffs	Empirical outcome*	Predicted expected total payoffs	Empirical outcome*	Predicted expected payoffs	Empirical outcome*
FP (85.56)	FP (80.98, 28.53)	FP (50.28)	FP (42.16, 15.65)	TD (42.78)	NoC (39.33, 12.52)
TD (81.30)	TD (79.21, 27.06)	TD (38.52)	TD (41.42, 15.36)	TDC (40.14)	FP (38.83, 12.53)]
TDC (65.28)	TDC (78.28, 27.41)	TDC (25.14)	TDC (41.04, 14.67)	NoC (37.5)	TD (37.79, 11.13)
NoC (60)	NoC (70.77, 33.26)	NoC (22.5)	NoC (31.44, 19.11)	FP (35.28)	TDC (37.25, 12.39)

* The numbers in columns labeled "empirical outcomes" are average expected period earnings and their respective standard deviations (in points). They are calculated based on the players' types (white or blue), the signals that the computer reported to the whites, the conditional probabilities of the states of the world (each signal is true with 70% probability) and the actual votes in a given period. In case of a voting tie, the expected earnings are based on the probabilities with which the policies are implemented ($\frac{1}{3}$ in case there is a tie between three policies, $\frac{1}{2}$ in case there is a tie between two policies). Treatments are abbreviated as follows: *FullyPublic*: *FP*, *TopDown*: *TD*, *TopDownClosed*: *TDC*, *NoChat*: *NoC*.

Table B.4: Expected period earnings across treatments

	(1) All Players	(2) Whites	(3) Blues	(4) All Players
FullyPublic (FP)	1.702***	3.572***	−0.168	−0.168
	(0.001)	(0.002)	(0.560)	(0.560)
TopDown (TD)	1.407***	3.327***	−0.512**	−0.512**
	(0.003)	(0.002)	(0.035)	(0.035)
TopDownClosed (TDC)	1.252**	3.199***	−0.694**	−0.694**
	(0.011)	(0.003)	(0.025)	(0.025)
Period	−0.036*	−0.104***	0.032	−0.036*
	(0.100)	(0.003)	(0.244)	(0.100)
White player				−2.629**
				(0.020)
FullyPublic × White player				3.740***
				(0.006)
TopDown × White player				3.839***
				(0.003)
TopDownClosed × White player				3.893***
				(0.002)
Constant	12.172***	11.573***	12.771***	13.486***
	(0.000)	(0.000)	(0.000)	(0.000)
Wald-test results for comparison of treatment coefficients (p values):				
FP vs. TD	0.347	0.709	0.154	
TD vs. TDC	0.607	0.800	0.446	
FP vs. TDC	0.213	0.557	0.085	
R^2	0.020	0.080	0.007	0.049
Number of clusters	20	20	20	20
Observations	9360	4680	4680	9360

Pooled OLS regressions. Dependent variable: Expected earnings (in points), conditional on received signals. Robust standard errors are clustered at the session level and p-values are given in parentheses: * p<0.10, ** p<0.05, *** p<0.01. NoChat serves as baseline treatment in all regressions.

Table B.5: Summary of hypotheses testing

Hypothesis	Behavior consistent with the hypothesis	Behavior inconsistent with the hypothesis
1. Truth-telling	More frequent in *FP* and *TD* than in *TDC*.	Truth-telling more frequent in *TD* than in *FP*, no difference between *FP* and *TDC*.
2. Blues' trustfulness	Blues more trustful in *FP* and *TD* than in *TDC*.	Blues not more trustful in *TD* than in *TDC*.
3. Whites' voting decisions	Given majority $x(y)$, whites' votes for $A(B)$ not different between *FP*, *TD* and *TDC* and higher than in NoC.	–
4. Blues' voting decisions	Given majority signal x, more blues vote for A higher in *FP* and *TD* than in *TDC* and *NoC*. Given majority signal y, more blues vote for B higher in *FP* than *TD*, *TDC* and *NoC*.	Given majority signal x, blues vote for A similarly often in *TD* and *TDC*.
Further predictions based on our theoretical results:		
5. Voting outcomes given majority signal x	*AA* more frequent in *FP* and *TD* than in *TDC* and *NoC*. *AC* weakly more frequent in *TDC* than in *NoC* and more frequent in *NoC* than in *FP* and *TD*.	*AA* not more frequent in *TD* than in *TDC* and *AC* more frequent in *NoC* than in *TDC* and not more frequent in *TDC* than in *FP* or *TD*.
6. Voting outcomes given majority signal y	*BB* more frequent in *FP* than in all other treatments. *BC* more frequent in *TD* and *TDC* than in *FP* and weakly more frequent than in *NoC*.	*BC* not more frequent in *TD* or *TDC* than in *FP*.
7. Efficiency	Higher in *FP* than in *TD*, in *TD* higher than in *TDC* and in *TDC* higher than in *NoC*. Efficiency in all communication treatment higher than in NoC.	Efficiency ranking of the communication treatments as predicted, but differences are not significant.

Treatments are abbreviated as follows: *FullyPublic*: *FP*, *TopDown*: *TD*, *TopDown-Closed*: *TDC*, *NoChat*: *NoC*.

C Chat dimensions

Dimension 1 – General classification of chats

Table C.1: Rater 1: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... stress the public spirit	0.22	0.22	0.23
... stress the interests of the own color group	0.11	0.12	0.01
... suspect lying	0.03	0.01	0.01
... stress trust	0.01	0.01	0.02
... mention circumstances as justification for behavior	0.22	0.14	0.10
... mention circumstances as information	0.27	0.28	0.21
... stress hope or optimism	0.07	0.13	0.10

Table C.2: Rater 1: Fraction of groups in which the blues...

	FullyPublic
... stress the public spirit	0.09
... stress the interests of the own color group	0.13
... suspect lying	0.16
... stress trust	0.03
... mention circumstances as justification for behavior	0.13
... mention circumstances as information	0.23
... stress hope or optimism	0.04

Table C.3: Rater 2: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... stress the public spirit	0.25	0.24	0.27
... stress the interests of the own color group	0.08	0.03	0.00
... suspect lying	0.02	0.00	0.00
... stress trust	0.02	0.00	0.02
... mention circumstances as justification for behavior	0.13	0.16	0.03
... mention circumstances as information	0.16	0.18	0.07
... stress hope or optimism	0.07	0.16	0.03

Table C.4: Rater 2: Fraction of groups in which the blues...

	FullyPublic
... stress the public spirit	0.16
... stress the interests of the own color group	0.11
... suspect lying	0.18
... stress trust	0.05
... mention circumstances as justification for behavior	0.12
... mention circumstances as information	0.13
... stress hope or optimism	0.07

Dimension 2 – Recommendations

Table C.5: Rater 1: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... recommend voting for A	0.73	0.70	0.71
... recommend voting for B	0.41	0.49	0.29
... recommend voting for C	0.06	0.01	0.01
... recommend something different	0.04	0.01	0.06

Table C.6: Rater 1: Fraction of groups in which the blues...

	FullyPublic
... recommend voting for A	0.53
... recommend voting for B	0.30
... recommend voting for C	0.54
... recommend something different	0.04

Table C.7: Rater 2: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... recommend voting for A	0.73	0.70	0.71
... recommend voting for B	0.41	0.49	0.29
... recommend voting for C	0.08	0.02	0.01
... recommend something different	0.04	0.01	0.06

Table C.8: Rater 2: Fraction of groups in which the blues...

	FullyPublic
... recommend voting for A	0.55
... recommend voting for B	0.30
... recommend voting for C	0.54
... recommend something different	0.03

Dimension 3 – Addressees

Table C.9: Rater 1: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... address the speech to all	1.00	1.00	1.00
... address their speech to their own color group only	0.28	0.22	0.01
... address their speech to the other color group only	0.34	0.19	0.14
... directly address a speech to s.o. from own color group	0.13	0.14	0.06
... directly address a speech to s.o. from other color group	0.18	0.00	0.00

Table C.10: Rater 1: Fraction of groups in which the blues...

	FullyPublic
... address the speech to all	0.99
... address their speech to their own color group only	0.37
... address their speech to the other color group only	0.38
... directly address a speech to s.o. from own color group	0.15
... directly address a speech to s.o. from other color group	0.21

Table C.11: Rater 2: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... address the speech to all	1.00	1.00	1.00
... address their speech to their own color group only	0.20	0.01	0.00
... address their speech to the other color group only	0.22	0.06	0.03
... directly address a speech to s.o. from own color group	0.09	0.04	0.02
... directly address a speech to s.o. from other color group	0.11	0.01	0.00

Table C.12: Rater 2: Fraction of groups in which the blues...

	FullyPublic
... address the speech to all	0.99
... address the speech to their own color group only	0.29
... address the speech to the other color group only	0.18
... directly address a speech to s.o. from own color group	0.09
... directly address a speech to s.o. from other color group	0.11

Dimension 4 – Showing respect

Table C.13: Rater 1: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... show respect	0.17	0.17	0.15
... behave disrespectfully	0.31	0.14	0.04
... behave neutrally	1.00	1.00	1.00

Table C.14: Rater 1: Fraction of groups in which the blues...

	FullyPublic
... show respect	0.13
... behave disrespectfully	0.28
... behave neutrally	1.00

Table C.15: Rater 2: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... show respect	0.06	0.01	0.03
... behave disrespectfully	0.31	0.02	0.01
... behave neutrally	1.00	1.00	1.00

Table C.16: Rater 2: Fraction of groups in which the blues...

	FullyPublic
... show respect	0.06
... behave disrespectfully	0.27
... behave neutrally	1.00

Dimension 5 – Specific recommendations I

Table C.17: Rater 1: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... recommend LTED	0.00	0.00	0.00
... recommend AC	0.02	0.00	0.00
... recommend AA	0.72	0.69	0.54
... recommend BB	0.40	0.46	0.26
... recommend CC	0.05	0.01	0.01
... recommend voting acc. to majority signal	0.14	0.14	0.43
... do not give any such recommendation	0.04	0.01	0.06

Table C.18: Rater 1: Fraction of groups in which the blues...

	FullyPublic
... recommend LTED	0.00
... recommend AC	0.07
... recommend AA	0.52
... recommend BB	0.27
... recommend CC	0.53
... recommend voting acc. to maj. signal	0.06
... do not give any such recommendation	0.03

Table C.19: Rater 2: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... recommend LTED	0.00	0.00	0.00
... recommend AC	0.02	0.00	0.00
... recommend AA	0.70	0.70	0.71
... recommend BB	0.40	0.49	0.29
... recommend CC	0.06	0.02	0.01
... recommend voting acc. to majority signal	0.03	0.00	0.00
... do not give any such recommendation	0.05	0.01	0.06

Table C.20: Rater 2: Fraction of groups in which the blues...

	FullyPublic
... recommend LTED	0.00
... recommend AC	0.06
... recommend AA	0.52
... recommend BB	0.29
... recommend CC	0.47
... recommend voting acc. to majority signal	0.01
... do not give any such recommendation	0.05

Dimension 6 – Specific recommendations II

Table C.21: Rater 1: Fraction of groups in which the whites give recommendations to both color groups...

	FullyPublic	TopDown	TopDownClosed
..., not mentioning signals or abstentions	0.95	0.98	0.75
..., mentioning signals, but not abstentions	0.08	0.14	0.43
..., not mentioning signals, but abstentions	0.01	0.00	0.01
... do not give any such recommendation	0.04	0.01	0.06

Table C.22: Rater 1: Fraction of groups in which the blues give recommendations to both color groups...

	FullyPublic
..., not mentioning signals or abstentions	0.93
..., mentioning signals, but not abstentions	0.03
..., not mentioning signals, but abstentions	0.01
... do not give any such recommendation	0.06

Table C.23: Rater 2: Fraction of groups in which the whites give recommendations to both color groups...

	FullyPublic	TopDown	TopDownClosed
..., not mentioning signals or abstentions	0.94	0.86	0.85
..., mentioning signals, but not abstentions	0.16	0.42	0.17
..., not mentioning signals, but abstentions	0.00	0.00	0.00
... do not give any such recommendation	0.05	0.01	0.06

Table C.24: Rater 2: Fraction of groups in which the blues give recommendations to both color groups...

	FullyPublic
..., not mentioning signals or abstentions	0.94
..., mentioning signals, but not abstentions	0.08
..., not mentioning signals, but abstentions	0.00
... do not give any such recommendation	0.05

Dimension 7 – Inter-group fairness and efficiency

Table C.25: Rater 1: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... mention relative payoffs (W vs. B)	0.10	0.09	0.09
... mention joint payoffs (W + B)	0.19	0.20	0.13

Table C.26: Rater 1: Fraction of groups in which the blues...

	FullyPublic
... mention relative payoffs (W vs. B)	0.11
... mention joint payoffs (W + B)	0.04

Table C.27: Rater 2: Fraction of groups in which the whites...

	FullyPublic	TopDown	TopDownClosed
... mention relative payoffs (W vs. B)	0.08	0.10	0.03
... mention joint payoffs (W + B)	0.13	0.23	0.26

Table C.28: Rater 2: Fraction of groups in which the blues...

	FullyPublic
... mention relative payoffs (W vs. B)	0.09
... mention joint payoffs (W + B)	0.04

D Translated instructions

Welcome to today's experiment!

You are taking part in a decision situation and it is possible for you to earn some money. The amount of money that you are able to win depends on your decisions and on the decisions of the other participants that are assigned to you. Moreover, it is influenced by the role that is randomly allocated to you. After having finished the experiment, we would like to ask you to fill in a short questionnaire.

Please note that from now on and throughout the experiment it is **not allowed to communicate** unless the computer explicitly asks you to do so. If you have any questions, please raise your hand out of your cubicle. One of the experimenters will come to you then. Throughout the experiment, it is forbidden to use mobile phones, smartphones, tablets or the like. Any violation of the rules leads to exclusion from the experiment and payment. All decisions are made anonymously, i.e. none of the participants learns about the identity of the others. Also the payment will be made anonymously at the end of the experiment.

Instructions

1. What's it about – An overview

[NoChat:] This experiment is about making a decision within a group between three different options A, B and C by way of vote.

[Deliberation / TopDown / TopDownClosed:] This experiment is about making a decision within a group between three different options A, B and C through communication and by way of vote.

A group consists of three „white" and three „blue" members. Your payment depends on the decision that the group makes regarding the possible options. It depends, first, on the fact which of the options will be implemented. Second, it is determined by the role you are assigned to – the "white" one or the "blue" one. And third, it also depends on the situation that occurs – this can be either X or Y. The graph below, comprising two tables, shows how many points a white and blue group member can earn given the three options and depending the situation that occurs – X (left table) or Y (right table).

	Situation X		**Situation Y**	
Options	White members	Blue members	White members	Blue members
A	20	20	10	0
B	0	0	20	10
C	0	10	0	20

The following applies for situation X: If option A is implemented, the white members and the blue members earn 20 points; if option B is implemented none of the members earns anything. If option C is implemented, the white members do not earn anything and the blue members earn 10 points.

The analogue applies for situation Y: If option A is implemented, the white members earn 10 points and the blue members do not earn anything; if option B is implemented, the white members earn 20 points and the blue members earn 10 points. If option C is implemented, the white members do not earn anything and the blue members earn 20 points.

The situation is not directly observable, but is selected randomly by the computer; both situations X and Y are equally likely i.e. they will be realized with a probability of 50%. The situation that is chosen by the computer is valid for the entire group; i.e. the payments for the white members as well as for the blue members are determined by either the left table or the right table. Thus, one could also say that the computer selects randomly one out of the two tables for the entire group, whereby both tables are equally likely.

Besides the partly different payments, there is also another difference between the white members and the blue members within a group: Each white member receives independent information by the computer on whether situation X or situation Y occurs. This information is true with a 70% probability (i.e. it is true in 70 out of 100 cases and wrong in 30 out of 100 cases). Thus, as this information does not always have to be true, it is possible that not all three white members receive the same information by the computer. The blue members do not receive any information by the computer.

[NoChat:] In order to make a decision between the three options, the group goes through a two-stage process. On the first stage, all group members can take notes in order to sort out their thoughts. On the second stage the voting will be carried out. The option with the most votes will be implemented.

[Deliberation:] In order to make a decision between the three options, the group goes through a two-stage process. On the first stage, all group members can chat together. On the second stage the voting will be carried out. The option with the most votes will be implemented.

[TopDown:] In order to make a decision between the three options, the group goes through a two-stage process. On the first stage, all white group members can chat together and send messages to the entire group. The blue members can read these messages, but they cannot actively take part in chatting. On the second stage the voting will be carried out. The option with the most votes will be implemented.

[TopDownClosed:] In order to make a decision between the three options, the group goes through a three-stage process. On the first stage, all white group members can chat together. The blue members cannot read these messages. On the second stage, all white group members can chat together and send messages to the entire group. The blue members can read these messages, but

they cannot actively take part in chatting. On the third stage the voting will be carried out. The option with the most votes will be implemented.

The experiment comprises 20 rounds.

In the following, the experiment will be explained in detail:

1. ***The allocation of the roles***

At the beginning of the experiment, the computer randomly assigns every participant either the role of a white member or that of a blue member. The **roles** remain **constant throughout the whole experiment**, i.e. one's own role will not change between rounds. Instead, in each round the group constellation will be re-determined: In each round the computer randomly allocates the participants to groups of six, consisting of three white members and three blue members.

In the following the course of an (arbitrary) round will be described. The experiment consists of **20 rounds**. The payments in any given round only depend on what happens in that round – they are independent of former rounds. The situation that occurs in a given round is likewise independent of the situations that have occurred in former rounds.

2. ***Course of a round***

At the beginning of each round, the computer randomly assigns the whites and the blues to groups of six, consisting of three white members and three blue members. Then each **white member** receives **information** by the computer on whether situation X or Y prevails, i.e. if the left or right table is correct. This information is true with a 70% probability. The blue members do not get any information.

[NoChat:] Then a **"note"**-window opens where you can write down notes. Please use the window only for taking notes regarding things that are relevant for the experiment. The window will disappear after **two minutes**. You will see in the top right corner how much time you have left.

[Deliberation:] Then a **"chat"**- window opens where **all group members**, the white and the blue members, can chat together. The computer randomly assigns everyone who enters a message a number that will be shown at the beginning of the message sent together with the role (white or blue). A possible pseudonym is for example „Blue 1". Please note: The pseudonyms are only valid **for this round**. With the help of these pseudonyms you can address each other and keep track of which messages are sent from the same person during the chat. Throughout the chat you can try to influence the voting decisions of the others. Please only use this chat for exchanging views on things that are relevant for the experiment. It is not allowed to uncover one's own identity or the identity of other group members. The chat window will disappear after **two minutes**. You will see in the top right corner how much time you have left.

[TopDown:] Then a "**chat"**- window opens where the **white group members** can chat together and send messages to the entire group. The blue members can read these messages, but they cannot

actively take part in chatting. The computer randomly assigns all white members who enter a message a number that will be shown at the beginning of the messages sent. A possible pseudonym is for example „White 1“.Please note: The pseudonyms are only valid **for this round**. With the help of these pseudonyms you can address each other and keep track of which messages are sent from the same person during the chat. Throughout the chat you can try to influence the voting decisions by the others. Please only use this chat for exchanging views on things that are relevant for the experiment. It is not allowed to uncover one's own identity or the identity of other group members. The chat window will disappear after **two minutes**. You will see in the top right corner how much time you have left.

[TopDownClosed:] Then a **"chat"**- window opens for the white group members where they can chat together. The blue members cannot read these messages. They have to wait for the experiment to proceed. Subsequently, another chat window opens where the white group members can chat together and send messages to the entire group. The blue members can read these messages, but they cannot actively take part in chatting. In both chats, the computer assigns all white members who enter a message randomly a number that will be shown at the beginning of the messages sent. A possible pseudonym is for example „White 1“. Please notice: The pseudonyms are only valid **for this round**. With the help of these pseudonyms you can address each other and keep track of which messages are sent from the same person during the chat. Throughout the chat you can try to influence the voting decisions by the others. Please only use this chat for exchanging views on things that are relevant for the experiment. It is not allowed to uncover one's own identity or the identity of other group members. Each of these chat windows will disappear after **one minute**. You will see in the top right corner how much time you have left.

[NoChat:] In the next step there is a secret **vote over the three options**. That means each group member can vote anonymously either for A or B or C or abstain from voting. Ultimately, the computer implements the **option with the most votes**. (In case of parity of votes the computer randomly chooses between the options with the most votes. Also in case that all group members abstain from voting, the computer randomly chooses one of the three options.)

[Deliberation / TopDown:] After the chat there is a secret **vote over the three options**. That means, each group member can vote anonymously either for A or B or C or abstain from voting. Ultimately, the computer implements the **option with the most votes**. (In case of parity of votes the computer randomly chooses between the options with the most votes. Also in case that all group members abstain from voting, the computer randomly chooses one of the three options.)

[TopDownClosed:] After the second chat there is a secret **vote over the three options**. That means, each group member can vote either for A or B or C or abstain from voting. Ultimately, the computer implements the **option with the most votes**. (In case of parity of votes the computer randomly chooses between the options with the most votes. Also in case that all group members abstain from voting, the computer chooses one of the three options.)

Then all group members are informed about the option that has been elected and they learn about the distribution of votes, i.e. how many votes option A has received, how many votes option B has received, how many votes option C has received and how many abstentions there have been. Moreover, the computer screen informs each group member about the situation that has occurred and how many points he or she has earned in the given round.

3. Total payment for the experiment

At the end of the experiment the computer will randomly, and independently from each other, selected three rounds. All rounds are equally likely. The payments that you have earned in these selected rounds will be summed up and converted into EURO with the **exchange rate 1 EURO = 3 POINTS**. Your total earnings from the experiment consist of the resulting amount plus the show-up fee of 10 EURO.