Supplementary Materials

EC.1. Simulation

EC.1.1. Network generation & selection

There are 3 networks of interest: Erdös-Rényi (ER), Stochastic Block (SB), and Royal Family (RF) network. All networks have 40 nodes, $n = 40$. In order to control for the average information received of each node, the networks have an average outdegree of 4 (excluding self links).

The generation process of each network type is as follows. The parameter specifications in the network generation process were selected to ensure strong connectedness in the networks generated.

- Erdös-Rényi networks are generated according to the Erdös-Rényi model (using the "erdos.renyi.game" function from the igraph package). We specify the number of nodes as n and total number of edges as $2n$.
- Stochastic Block networks are generated according to the Trait-based random generation (using "sample_pref" function from the igraph package). We specify the number of nodes as n and the size of each community as 5. So there are $n/5$ communities where the probability of linking within a community is $p_{ii} = 0.85$ and between communities is $p_{ij} = p_{ii}/60$. (These parameter specifications were selected to ensure strong connectedness in the networks generated)
- Royal Family networks are created by first placing n players in a directed ring (player n observes player 1 who then observes player 2 and so on). Then players 1,2,3 are selected to be the hub where all players observe them. All players have outdegree of 4 (except for player 1 and 2 with outdegree of 2, and player 3 and n with outdegree of 3).

For each network treatment, we randomly generated 100 networks that are (strongly) connected — every node can be reached through a path from every other node. Then we computed network measures such as outdegree, diameter, average path length, clustering for each network. The average statistics for each network type are presented in Table EC.1.

Out of the 100 randomly generated networks, a network with measures closest to the average statistics is then used in the experiment. Table EC.2 presents the network statistics of these networks and Figure EC.1 presents the network graphs. Note that the Royal Family network is not generated randomly.

EC.1.2. Signal generation & selection

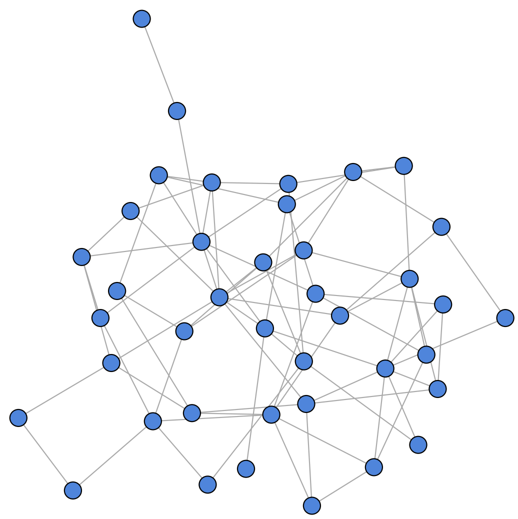
We randomly select 24 sets of signals for the experiment. For each network treatment, there are 4 groups of players each playing 6 rounds. So group 1 in round 1 uses the first set of signals while

Table EC.1 Averages network statistics of 100 randomly generated networks

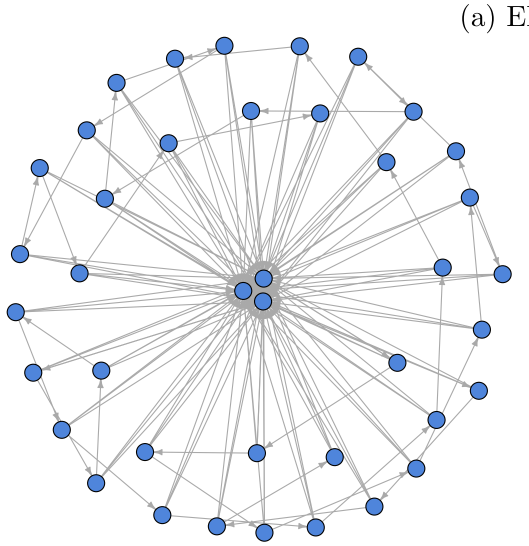
n=40	avg. outdegree	diameter	avg path length	clustering
ER	4.00	5.63	2.73	0.10
SB	3.98	9.15	4.12	0.57
RF	3.85	38.00	12.72	0.26

Table EC.2 Network statistics of the networks used in the experiments

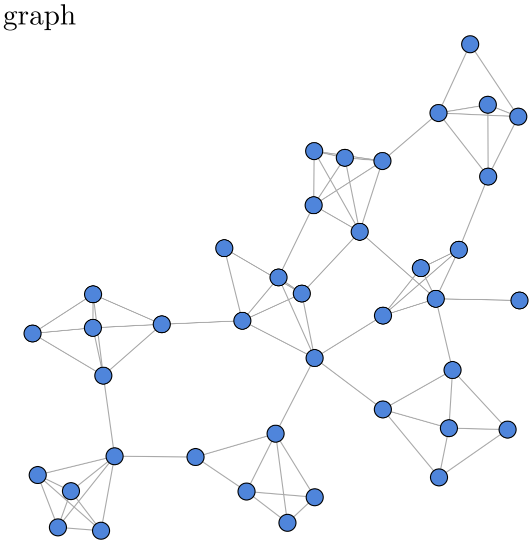
n=40	avg. outdegree	diameter	avg path length	clustering
ER	4.00	5	2.73	0.10
SB	4.00	9	3.85	0.57
RF	3.85	38	12.72	0.26



(a) ER graph



(b) RF graph

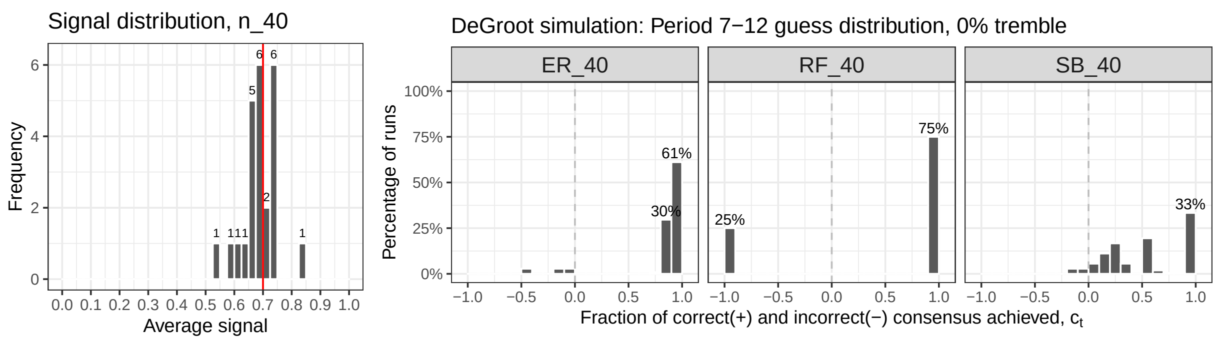


(c) SB graph

Figure EC.1 Network graph of n=40 with average outdegree 4

group 4 in round 6 uses the 24th set. Therefore, the same collection of signals are used across all networks.

We perform two checks to ensure that the 24 sets of signals are representative. First, we note that the distribution of the 24 sets of signals is bell-shaped around the mean 0.7 where 1 represents the correct state (Figure EC.2a). Second, we confirm that the simulated guesses following these 24 sets of signals (Figure EC.2b) have the same properties as the simulations of the 1000 sets of signals (presented in Figure 2c, see main text). The regression on network effects with respect to 'Correct consensus', 'Incorrect consensus' and 'Breakdown of consensus' (as defined in the Consensus Outcomes section in the main text) confirms the main hypotheses (Table EC.3).



(a) Distribution of signals (b) Distribution of c_t simulated from signals

Figure EC.2 Signal distribution and simulation results using the 24 sets of signals for the experiment. (a) Distribution of signals used for all networks in the experiment with mean 0.70, standard deviation 0.06, 1st quartile 0.675, 2nd quartile 0.70 and 3rd quartile 0.75. (n=24) (b) Distribution of c_t under DeGroot simulation using experiment signals. The hypotheses from the simulation of 1000 runs are confirmed: 1) There is more breakdown of consensus in the Stochastic Block network than in the Erdös-Rényi and Royal Family network; 2) There is more incorrect consensus in the Royal Family network than in the Erdös-Rényi and Stochastic Block network.

Table EC.3 OLS regression of simulated data, network size 40, k = 0.3

	OLS - Correct Consensus	OLS - Incorrect Consensus	OLS - Breakdown
(Intercept)	0.88^{***}	0.04	0.08
	(0.09)	(0.05)	(0.07)
typeRF	-0.08	0.17^{**}	-0.08
	(0.12)	(0.08)	(0.10)
typeSB	-0.37^{***}	-0.04	0.42^{***}
	(0.12)	(0.08)	(0.10)
R^2	0.13	0.11	0.31
Adj. R^2	0.10	0.08	0.29
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

EC.1.3. Variations on DeGroot updating rule

All network effects identified in the simulations are robust to alternative variations on DeGroot updating rule.

Deterministic DeGroot. In the case of indifference, suppose an individual persists with her last period's guess. Formally, we say:

$$a_{i,t} = \begin{cases} 1 & \text{if } \mu_{i,t} > \frac{1}{2}, \\ 0 & \text{if } \mu_{i,t} < \frac{1}{2}, \\ a_{i,t-1} & \text{if } \mu_{i,t} = \frac{1}{2} \end{cases} \tag{EC.1}$$

Simulations of this variant of the DeGroot are presented in Figure EC.3.

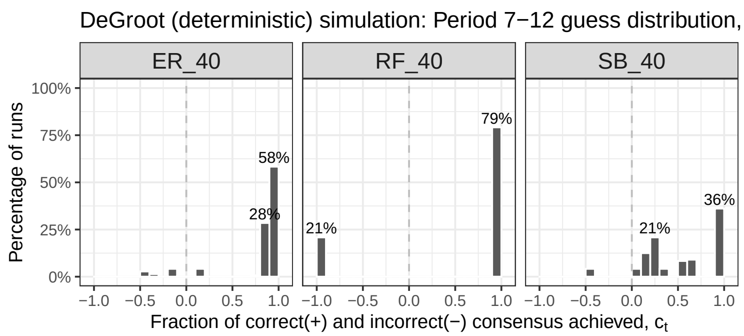
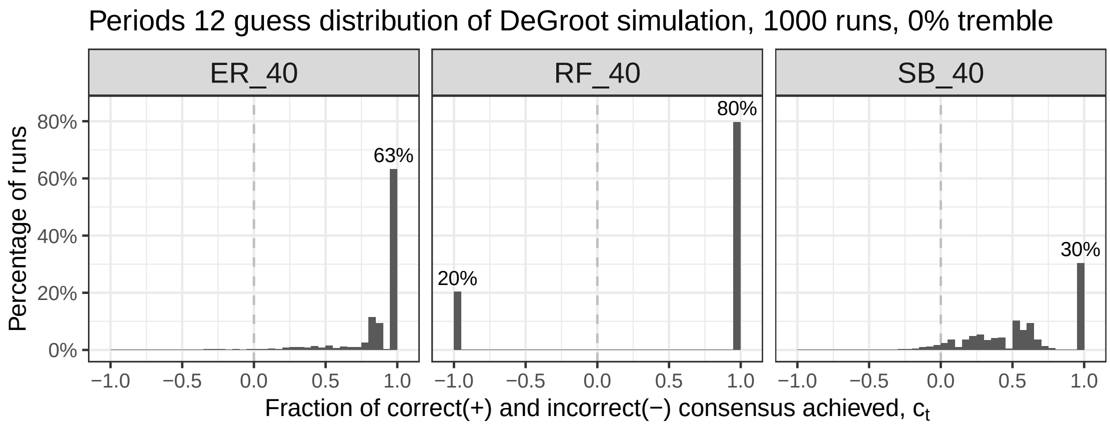


Figure EC.3 **Distribution of c_t under Deterministic DeGroot simulation using experiment signals.**

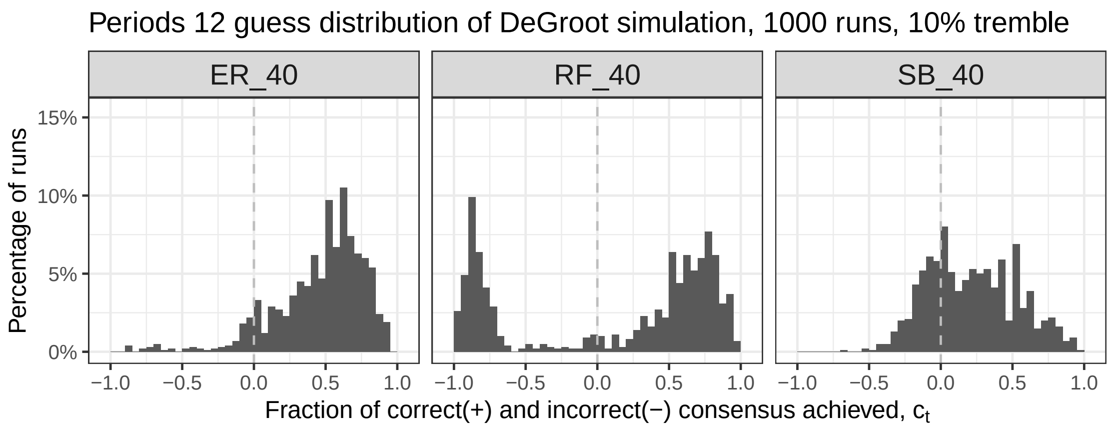
DeGroot with Trembling. Suppose an individual observes a majority guess of Red: if we use DeGroot updating rule with 10% trembling, that means she would guess Green 10% of the time and Red 90% of the time. Figure EC.4 shows that the networks effects identified with the original DeGroot (as in Figure 2c in the main text) are robust.

EC.1.4. Welfare

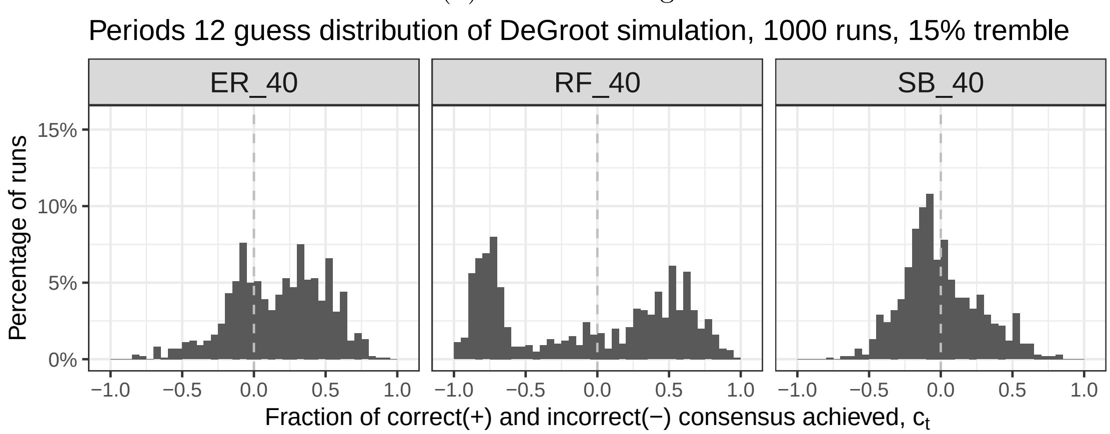
To compare welfare performances across networks, we use a simple measure of welfare improvement: $w_t = (n_t - n_0)/(n - n_0)$, i.e., the first part of c_t in eq. (2), applied to the entire range. The measure ranges from 1 (maximum gain in welfare) to $-n_0/(n - n_0)$ (maximum loss in welfare). The simulation shows that, in the limit, Erdös-Rényi has the highest average welfare improvement, followed by Stochastic Block, and Royal Family comes last (Figure EC.5). The reason for the poor showing of the Royal Family network is the high frequency of incorrect consensus outcomes. Stochastic Block network achieves lower welfare improvement than Erdös-Rényi network because it has more breakdowns of correct consensus.



(a) 0% trembling



(b) 10% trembling



(c) 15% trembling

Figure EC.4 Distribution of c_t under simulation with trembling.

EC.2. Findings

EC.2.1. Convergence

The rapid convergence of guesses in the experiment is supported by evidence on switching frequency: 20% of individuals switched their guesses at period 2 after observing the first period guesses

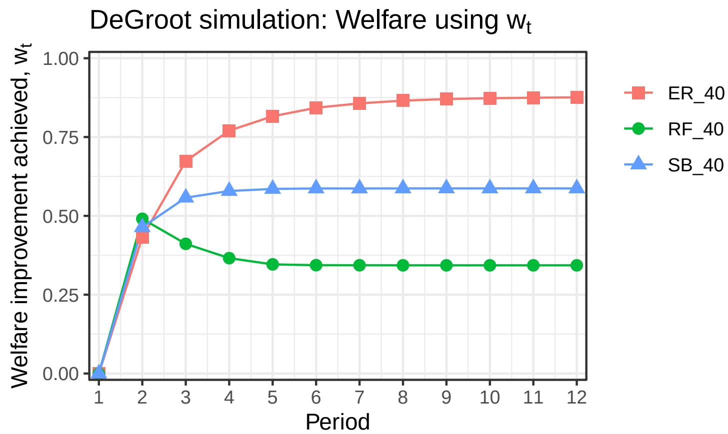
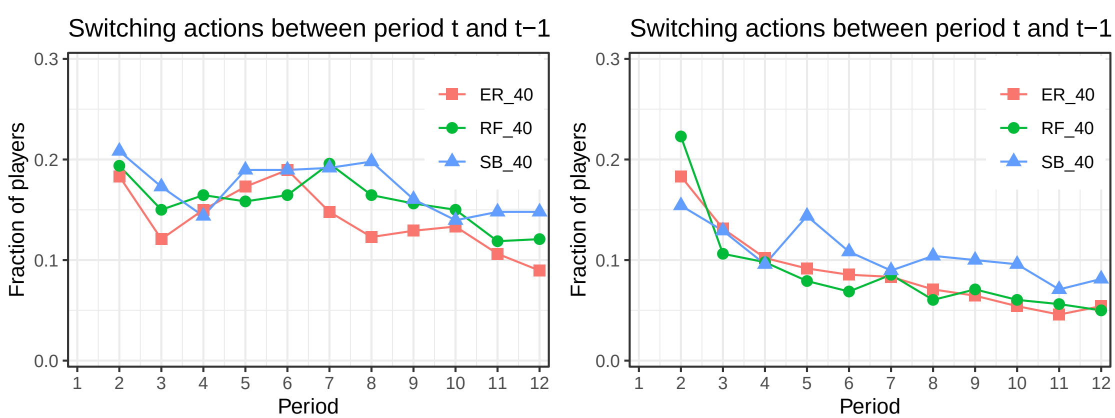


Figure EC.5 Evolution of welfare improvement w_t in the simulation: In period 12, ER achieved 88% of the possible welfare improvement, 58% for SB and 34% for RF.

of their neighbors, this switching frequency falls to 10% toward the end of the experiment in period 12. The switching probability falls significantly as subjects learn across rounds: as a result, it is only 5% in the last three rounds (Figure EC.6). We argue that the residual switching in guesses in the final periods are not due to further learning by subjects, but due to random guessing.



(a) n=40, round 1-3 (b) n=40, round 4-6

Figure EC.6 Percentage of subjects switching guesses per period. (a) In the first three rounds, percentage of switching falls from 20% in period 2 to 10∼15% in period 12. (b) In the last three rounds, percentage of switching falls from around 20% in period 2 to 5∼8% in period 12. Therefore, adjusting for learning across rounds, there is less than 8% of subjects switching guesses by period 12.

We estimate that 10% of the guesses are random in the experiment, using the following technique: Irrespective of whether a myopic player follows Bayesian or DeGroot learning rule, in period 1, it is optimal to guess her initial signal. In period 2, both (myopic) Bayesian and DeGroot learning rules predict that player should follow the majority guess in her neighbourhood in period 1. Table EC.4

shows that about 10% of guesses do not follow subjects' initial signals in period 1 and contradict both learning rules in period 2. This suggests that about 10% of guesses ignore information.

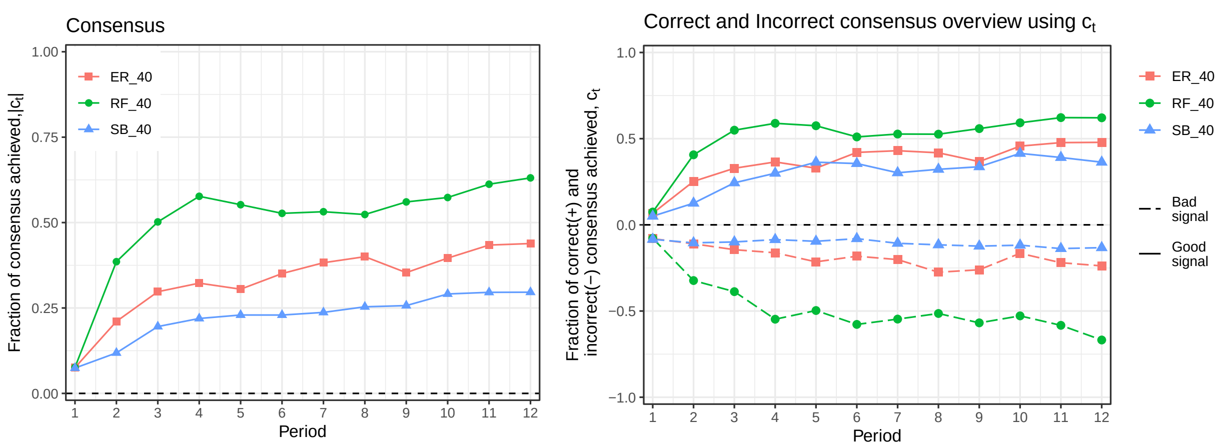
Table EC.4 Fraction of guesses against Bayesian and DeGroot prediction, network size 40

	Guess against majority in period 1,2	
	OLS (Bayesian, DeGroot predicts 0)	Logit
(Intercept)	0.10^{***}	-2.24^{***}
	(0.01)	(0.08)
typesizeRF_40	0.02^{*}	0.24^{*}
	(0.01)	(0.13)
typesizeSB_40	0.02^{***}	0.25^{***}
	(0.01)	(0.08)
R^2	0.00	
Adj. R^2	0.00	
Num. obs.	5760	5760
AIC		4029.57
BIC		4049.54
Log Likelihood		−2011.78
Deviance		4023.57

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

EC.2.2. Consensus

The simulations lead us to propose two hypotheses: One, the breakdown of consensus is most likely in the Stochastic Block network, followed by the Erdös-Rényi network and lastly the Royal Family network; Two, the Royal Family network leads to the wrong consensus more often than the Erdös-Rényi network. Figure EC.7a presents the evolution of consensus across periods across all networks, while Figure EC.7b presents the evolution of c_t partitioned by ‘good’ and ‘bad’ signals. Under DeGroot updating simulation, the set of ‘good’ signals would lead to $c_t \geq 0$ (correct consensus), while the ‘bad’ signals would lead to $c_t < 0$ (incorrect consensus). They show that the rankings in the hypotheses are maintained across all periods. The regression Table EC.7 shows the statistical significance of the estimates (presented in Figure 4 in the main text), supporting our hypotheses. The estimate of ‘incorrect consensus‘ on ‘typeRF’ represents the difference in fraction of incorrect consensus achieved between the Royal Family network and the Erdös-Rényi network.



(a) Evolution of consensus $|c_t|$

(b) Evolution of consensus c_t

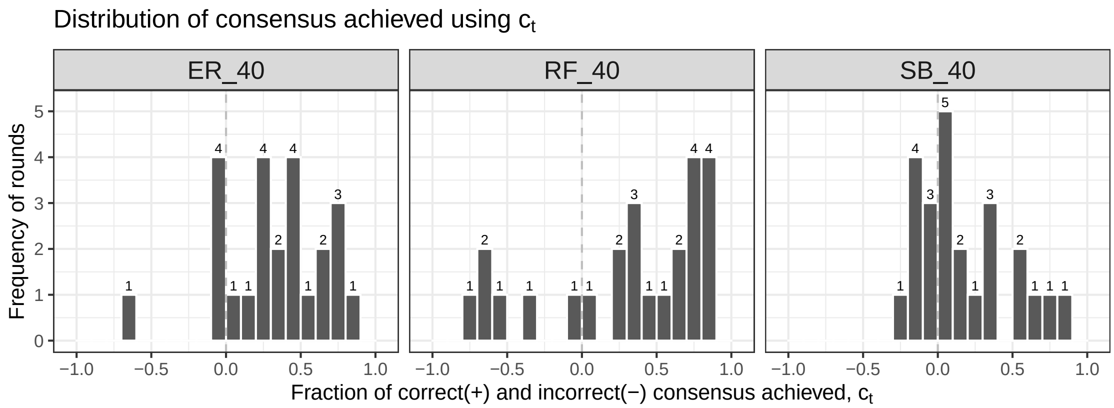
Figure EC.7 Evolution of $|c_t|$ and partitioned c_t. (a) In period 12, RF, ER, SB reach 63%, 44%, 30% of consensus, respectively. (b) We partitioned c_t averaged across all games by ‘good’ and ‘bad’ signals. The ranking of correct and incorrect consensus reached is preserved across most periods.

A similar distribution of c_t obtains if we consider fewer periods (periods 10-12) or rounds (rounds 4-6) (Figure EC.8).

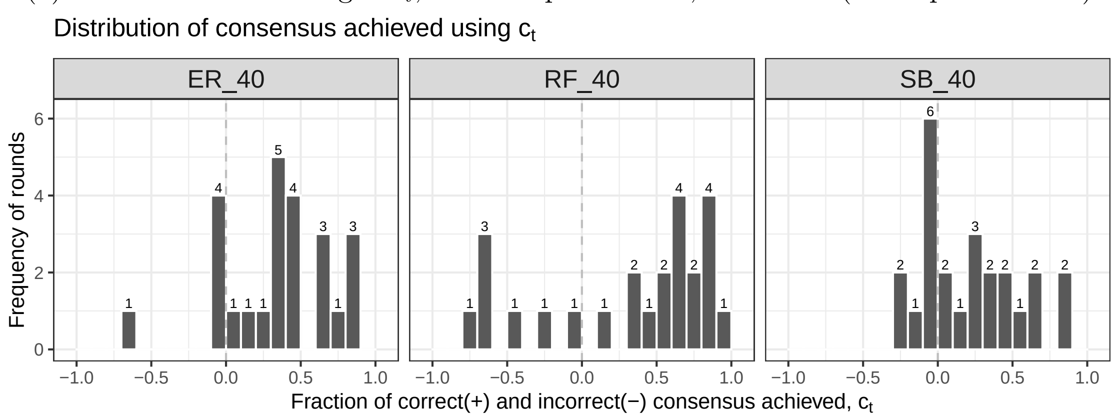
Recall, we defined binary variables of correct consensus (if $c_t > k$), incorrect consensus (if $c_t < -k$), and breakdown of consensus (if $-k \leq c_t \leq k$) based on the value of c_t. Our main findings are robust to 1) different widths k (Tables EC.6 to EC.8), 2) an alternative model specification such as the logit model (Table EC.9), and 3) a continuous definition of consensus outcomes (Table EC.10).

A continuous variation on the definition of consensus would be as follows: Consensus is defined as the absolute value of c_t, $|c_t|$; correct consensus is defined as censoring negative values of c_t to

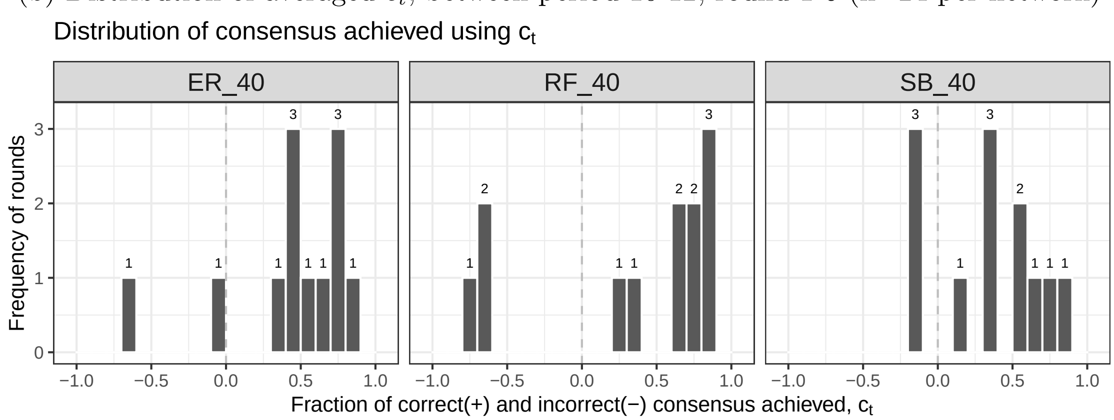
0; incorrect consensus censors positive values of c_t to 0; breakdown is defined as the negative of consensus, $-|c_t|$.



(a) Distribution of averaged c_t, between period 7-12, round 1-6 (n=24 per network)



(b) Distribution of averaged c_t, between period 10-12, round 1-6 (n=24 per network)



(c) Distribution of averaged c_t, between period 7-12, round 4-6 (n=12 per network)

Figure EC.8 Distribution of averaged c_t robust over period and round selections.

For the Stochastic Block network, we show that communities are more likely to reach consensus compared to Erdös-Rényi network despite the networks being less likely to reach consensus. The subset of subjects in each block of the Stochastic Block model may be seen as constituting a 'community'. Given the network generation methods in the Stochastic Block model, subjects with location id $1-5$ is a community while id $6-10$ is another community, and so on. So in all three networks, we define a community by the same location ids. We define *community consensus* as 1 when all 5 subjects in a community reaches complete consensus, and 0 otherwise.

Table EC.5 shows that 52% of communities reach consensus in the Stochastic Block network which is 14% point higher than in Erdös-Rényi network. We next look closer at the dispersion of average guesses of communities. The maximum difference in average guesses of communities is equal to 1: when there exists one community with correct consensus and one with incorrect consensus. Figure EC.9 shows that 75% of rounds in the Stochastic Block network have large dispersion in community guesses (greater than 0.7) while only 50% in Erdös-Rényi network and 46% in Royal Family network. This implies that disagreements between communities are the principal source of the consensus breakdown in the Stochastic Block network.

Table EC.5 Regression of community consensus on network treatment

	OLS - Community Consensus	Logit - Community Consensus
(Intercept)	0.38^{***}	-0.49^{***}
	(0.03)	(0.15)
typeRF	0.16^{*}	0.66^{*}
	(0.09)	(0.37)
typeSB	0.14^{***}	0.55^{***}
	(0.04)	(0.16)
R^2	0.02	
Adj. R^2	0.02	
Num. obs.	576	576
AIC		791.85
BIC		804.92
Log Likelihood		-392.93
Deviance		785.85

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

EC.2.3. Welfare

We examine the implications of learning outcomes on welfare by using the welfare improvement measure w_t. In line with the DeGroot simulation, the Erdös-Rényi network has higher welfare improvement than the Stochastic Block network (Table EC.11). We find that the Erdös-Rényi network achieves the average welfare improvement of 0.30, while only 0.18 for the Stochastic Block network. We do not find any significant difference in welfare improvement between the Erdös-Rényi network and the Royal Family network. This is probably due to the large variation in welfare improvement and deterioration among the Royal Family networks.

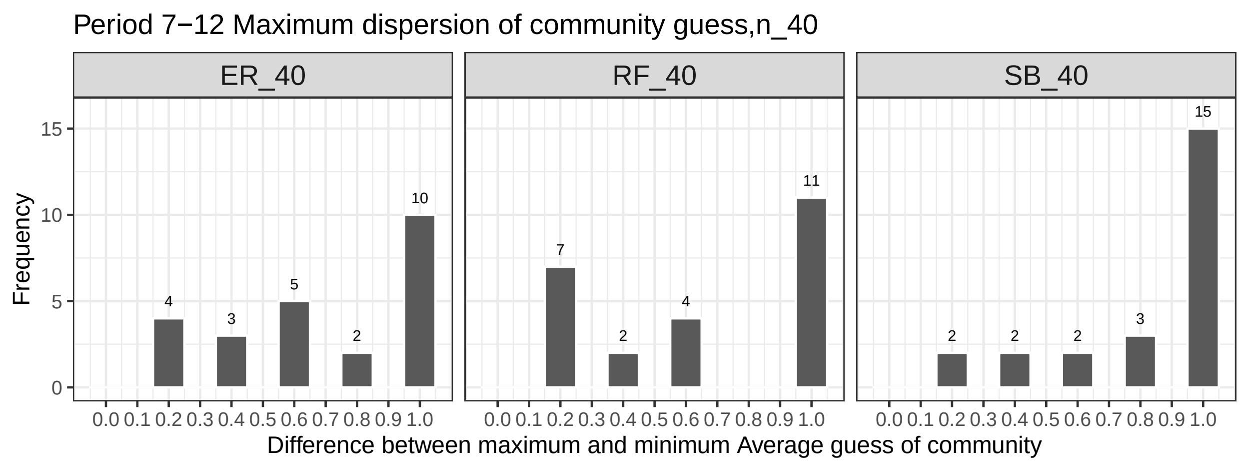


Figure EC.9 Distribution of the maximum dispersion in average guesses between communities for each network. 75% of rounds in the SB have more than 0.7 dispersion in average guesses between communities, 50% in ER and 46% in RF (n=24 per network).

EC.2.4. Updating rule

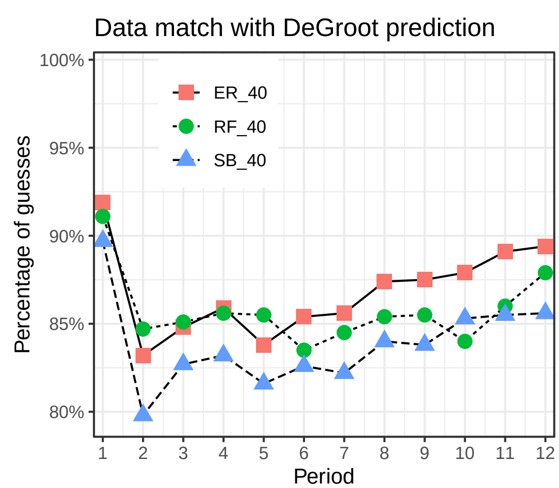
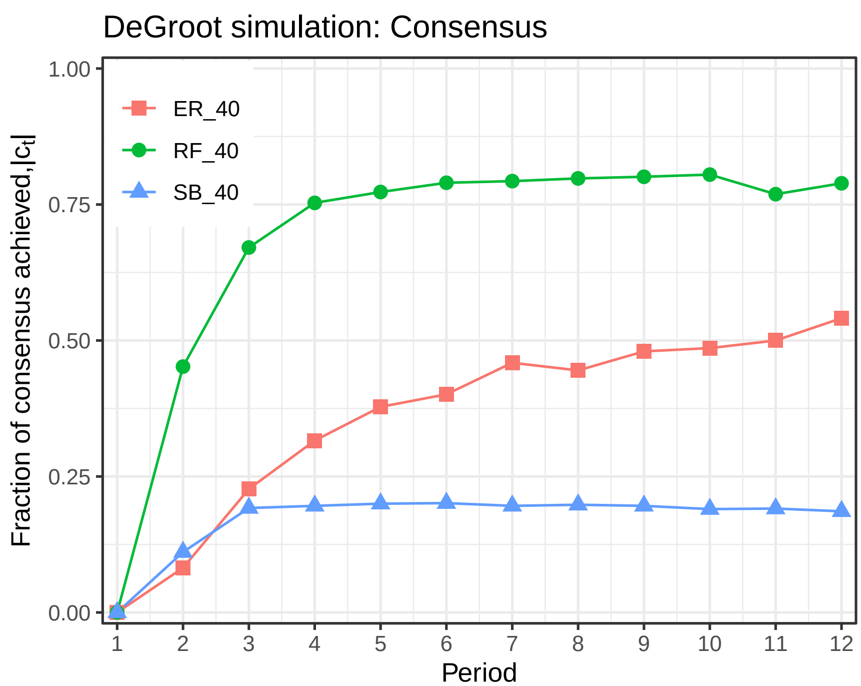


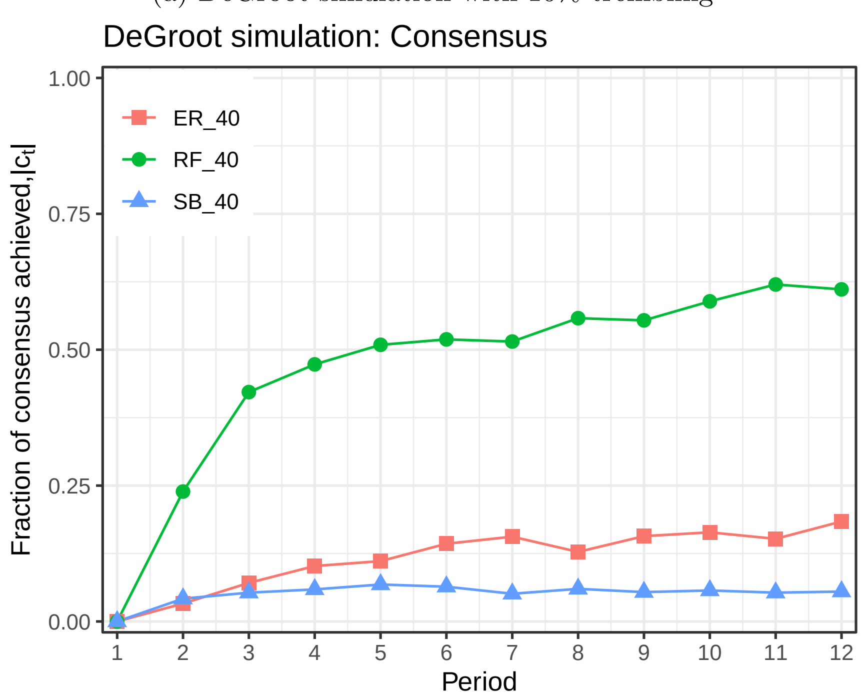
Figure EC.11 The percentage of guesses matching DeGroot prediction across periods. Across all networks at least 90% of first guesses matched with the DeGroot prediction (i.e., guess follows the signal). This percentage falls to 80%∼85% in the second period and then steadily increases until it reaches 85%∼90% in later periods.

On average, 88% of guesses match with the DeGroot rule. This is higher than the baseline of how well *guessing randomly* matches with DeGroot predictions: simulations show that on average 60% pseudo subjects' random guesses match with DeGroot. This is also higher than the baseline of how well *guessing signal* matches with DeGroot predictions: simulations show that on average 75% guesses of pseudo subjects (if guessing only signal) match with DeGroot (Figure EC.13a).

Suppose that 10% of guesses are randomly made. We show that the level of consensus attained in the experiment is comparable the simulation under 10% trembling for Erdös-Rényi (Figure EC.12a) and 15% trembling for Stochastic Block and Royal Family network (Figure EC.12b).



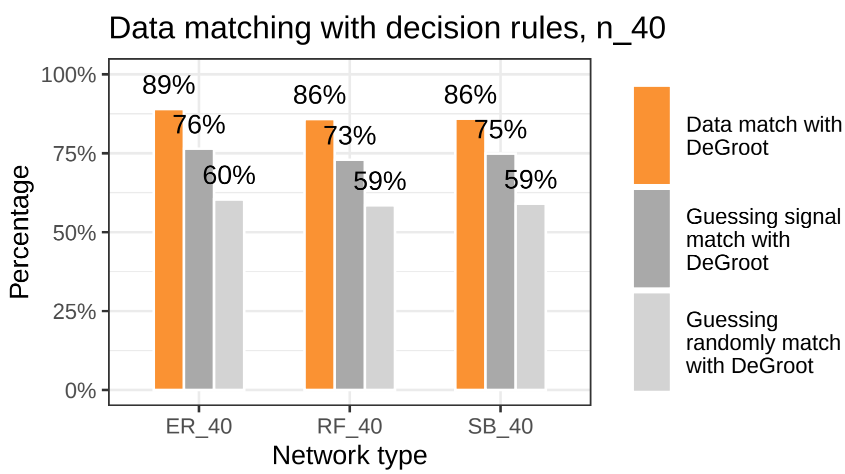
(a) DeGroot simulation with 10% trembling



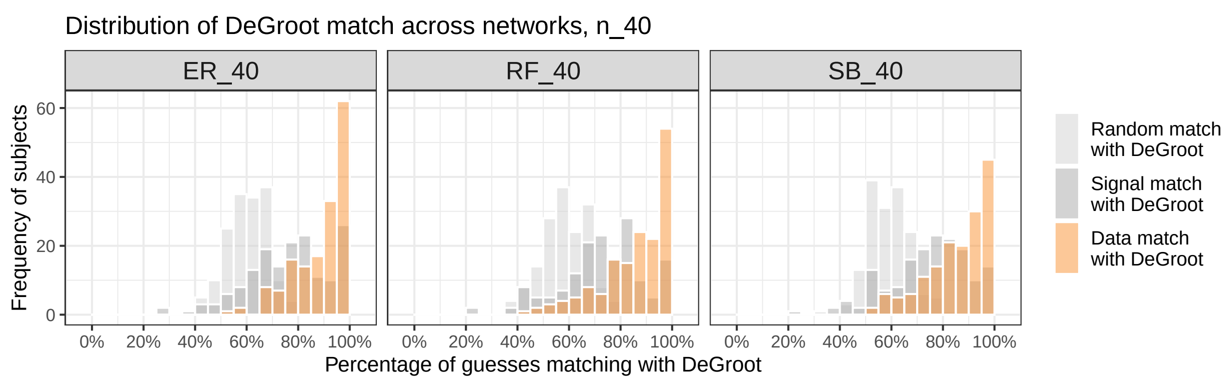
(b) DeGroot simulation with 15% trembling

Figure EC.12 Consensus achieved under DeGroot simulation with trembling.

We next delve deeper by looking at subject level match with DeGroot. Because each subject plays a total of 6 rounds and 12 periods per round and their guesses are not statistically independent, we treat each subject as a data point. Figure EC.13 presents the histogram of how well a subject's guesses match with DeGroot predictions. For all networks, there are significantly more subjects whose guesses match with DeGroot than pseudo subjects who guess their signals or randomly.



(a) Fraction of guesses that match with DeGroot, compared to baselines.



(b) Fraction of subjects that match with DeGroot, compared to baselines

Figure EC.13 Percentage of guesses/subjects match with the DeGroot rule. Guesses matching DeGroot prediction are in orange; (Simulation) Guessing signal matching DeGroot prediction are in dark grey; (Simulation) Guessing randomly matching DeGroot prediction are in light grey. (a) Roughly 88% of guesses match with DeGroot predictions, significantly higher than the other two baselines of 75% and 60% respectively. (n=46,080: 11,520 per network) (b) 80% of subjects in ER match with DeGroot predictions at least 80% of the time; these fractions are 72% in the RF and 76% in the SB. This is again compared to the baseline of how well guessing signal matches with DeGroot predictions: Only 44% of pseudo subjects' guesses (if guessing only signal) in ER match with DeGroot predictions at least 80% of the time (37% in RF, and 41% in SB); A negligible fraction of pseudo subjects' guesses (if guessing randomly) match with DeGroot predictions at least 80% of the time. (n=960: 240 per network)

Bayesian learning: Information Leader. When DeGroot prediction contradicts with information leader's guess, a Bayesian player should follow their information leader while a DeGroot player should follow the majority of their neighbours. Table EC.12 show that when the two are in conflict, only around 10% of subjects follow Bayesian prediction (ER:10%, RF:4%, SB:14%), while the rest follow DeGroot. Note that this percentage is decreasing in rounds which suggests subjects' learning behaviour towards DeGroot updating rule.

No learning: Stubborn players that only follow their signal. Similarly, when DeGroot prediction contradicts goes against initial signal received, a stubborn player should only follow their own signal. Table EC.13 show that around 25% of subjects follow initial signal (ER:25%, RF:29%, SB:29%) while the rest follow DeGroot. As before, we show that there is significant learning across rounds. More interestingly, increases in periods also decreases stubbornness. This could be due to increasing availability of information and therefore less weight is put on initial signals.

For robustness, we also replicate the results from Figure 5 by removing data points corresponding to indifference as predicted by the DeGroot decision rule (i.e., when $\mu_{i,t} = \frac{1}{2}$). This allows us to check that the agreement measures from Figure 5 are not inflated mechanically by those indifference data points. The corresponding results are depicted in Figure EC.14.

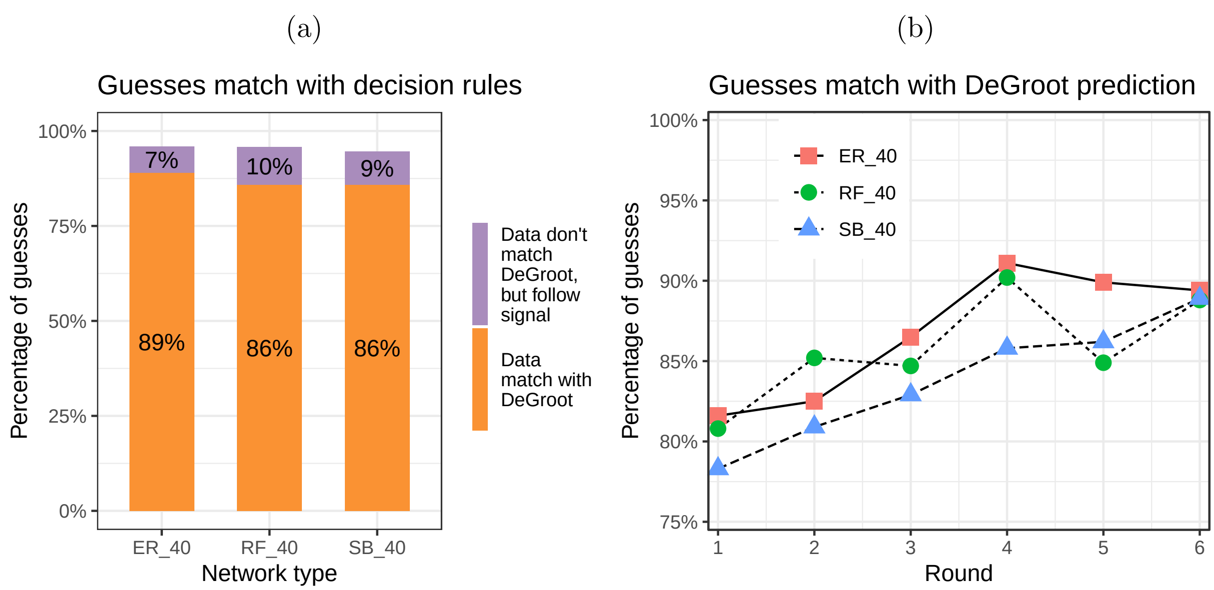


Figure EC.14 Comparing actual guesses with DeGroot prediction when removing data points where DeGroot predicts indifference. (a) 86∼89% of guesses match with DeGroot prediction and 7∼10% match with signal. Together they explain 95% of variation in guesses. (b) 78∼82% of guesses match with DeGroot prediction in round 1; this increases to 88∼90% by round 6 (n=33215: 10877 in ER, 11490 in RF, and 10848 in SB).

EC.2.5. Choice activity

We here present details about the subjects' behavior in any period of any round of the experiment. Figure EC.15 shows the timing of decisions within the 30 second constraint imposed to them. Results indicate that the vast majority of choices (87%) are made by second 10.

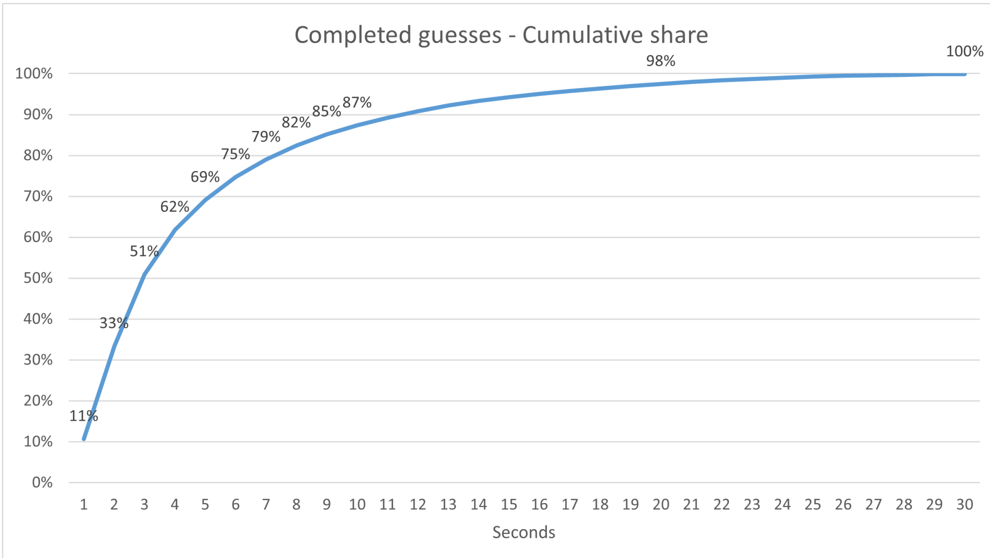


Figure EC.15 Overall cumulative distribution of decisions across time (in seconds) in any period of any round, and for any network. By second 5, 69% of guesses have been completed, and it reaches 98% by second 20 (subjects had 30 seconds to make each decision).

At any start of a new period of the game, subjects were shown the outcome of the previous period on the screen (through the network visualization tool). However, in any period t>2, they also had the opportunity to observe any earlier period of their choice. In Table EC.14a, we therefore check the frequency of subjects using this tool. Results show that a large fraction of subjects (45%) never chose to observe an earlier period. However, the likelihood of viewing an earlier outcome at any period is very low as highlighted by Table EC.14b.

As a means to identify the possible influence of such a history viewing tool on subject's behavior, we compare the frequency of their choices matching with the DeGroot prediction in every period, conditional on whether they observed some earlier period(s). Results presented in Table EC.15 reveal similar behavior regardless of whether they observed earlier periods (differences are not statistically significant). In other words, behavior is not influenced by their history viewing activity.

EC.3. Related experiments

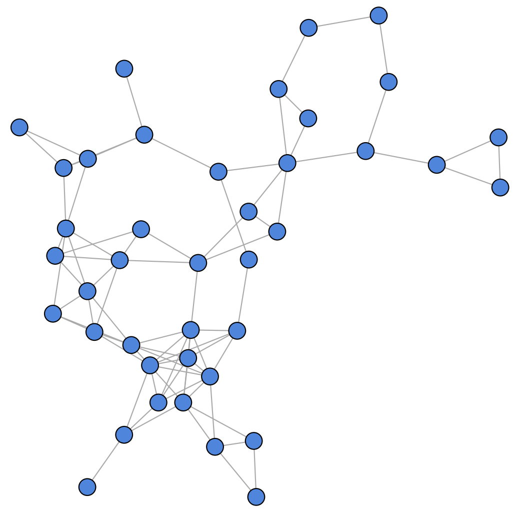


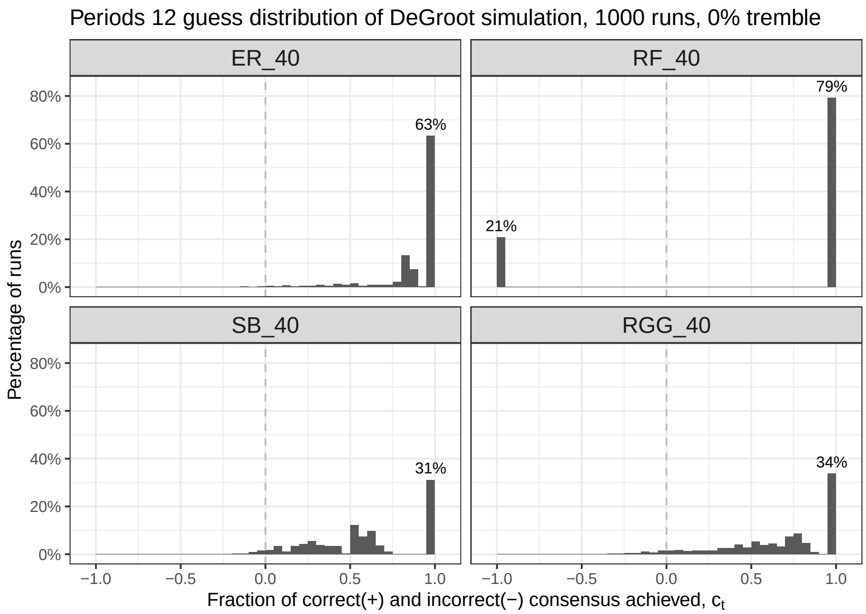
Figure EC.16 RGG graph of n=40 with average outdegree 4

In a recent paper, Chandrasekhar et al. (2020) looked at the mixture model of Random Geometric Graphs and Erdös-Rényi Graphs. We denote it as the RGG network from this point forward. This model captures the idea of sparse and clustered networks from the real world where the share of 'clans' — a set of nodes that are more connected among themselves than to those outside — is non-vanishing as n grows. This feature of inward-looking clans is also present in the 5-player communities within the Stochastic Block network. Under DeGroot updating rule, 'clans' being inward-looking facilitates the breakdown of consensus.

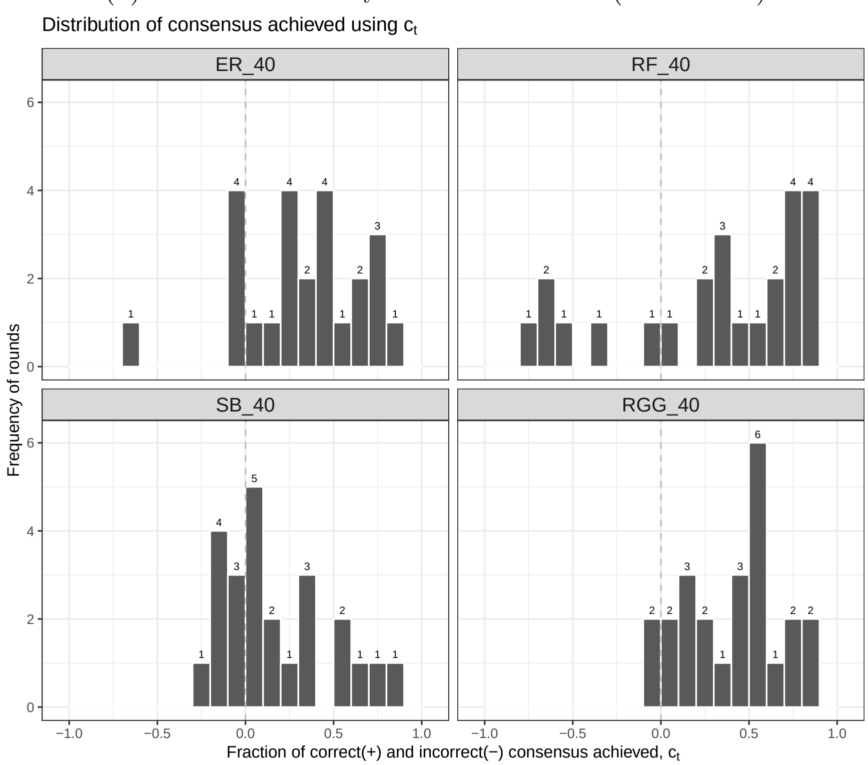
The network generation process is as follows: These exists a Poisson point process on the latent space $\Omega = [0,1]^2 \subset \mathbb{R}^2$, which determines the latent location of n nodes, with uniform intensity $\lambda > 0$. For any subset $A \subset \Omega$, $n_A \sim Poisson(\nu_a)$, where $\nu_a := \lambda \int_A dy$. If the Euclidean distance between two nodes i and j are at most $r = 0.2$, then i and j are linked with probability $\alpha = 0.95$. Otherwise, they are linked with probability $\beta = \alpha/(3n) < \alpha$. These parameter specifications were selected to ensure strong connectedness in the networks generated. Figure EC.16 presents an example of the RGG network which is also used in the experiment.

Figure EC.17a presents the simulation results of DeGroot updating rule on the RGG network and compares it with the Erdös-Rényi, Royal Family and Stochastic Block network. These simulations suggest that the RGG "lies between" the Erdös-Rényi and Stochastic Block network. The quartiles and the mean of the distribution of simulated c_t confirm this (Table EC.16). We observe the same results in the experiment (Figure EC.17b).

The Erdös-Rényi and Stochastic Block networks are canonical networks. Given the simulations and the experimental findings noted above, for expositional reasons, we felt it was best to present the Erdös-Rényi and Stochastic Block networks in the main text and move the RGG network to this Appendix.



(a) Distribution of c_t under Simulation (with RGG)



(b) Distribution of averaged c_t from experimental data (with RGG)

Figure EC.17 Distribution of averaged c_t. (a) The simulation of 1000 sets of signals shows that the distribution of consensus achieved by RGG lies between ER and SB. (b) Our experiment confirms the results from the simulation. (n=24 per network)

EC.4. Experimental Design

The experiment took place at the Laboratory for Research in Experimental and Behavioral Economics (LINEEX) at the University of Valencia. Subjects were recruited through the online recruitment system of LINEEX. All subjects who participated in this study provided informed consent at the LINEEX laboratory, and the procedure of this study was approved by the Institutional Review Board of the University of Valencia. In the experiment, subjects interacted through computer terminals in the LINEEX laboratory, and the experimental software was programmed in HTML, PHP, Javascript, and SQL.

Upon starting an experimental session, subjects read the paper-based instructions, which were also read aloud by an experimenter to guarantee that everyone received the same information (Supplementary Materials). The subjects were then provided with a step-by-step interactive tutorial on their computer screen, which allowed them to get familiarized with the software interface and the game (Figure EC.18). To clarify possible consequences of guesses in different periods of a round, subjects were shown a sample network (with only 10 players but with similar features as the network used in the actual game, depending on the experimental condition) highlighting what guesses would be observed by subjects as a decision maker from their neighbours, and their neighbours' neighbours.

Details about the decision screens were also provided to subjects: during any period of the game, each subject was shown the colour of the ball initially drawn, and guesses made by neighbours in the network during the previous period (Figure EC.20). Subjects also had the ability to view guesses made by those individuals (and themselves) in earlier periods of the game through a slider button. At the end of a round, a feedback screen revealed information about the payoff effective period that has been randomly selected, the guesses made by the subject and all others in thin period, the bag actually selected, and consequently the payoffs received by the subject in this round (0 or 3 euros depending on whether the guess matches the bag) (Figure EC.21). Prior to starting the first round of the game, all subjects also filled up a short questionnaire (4 questions) about their comprehension of the decision screens (Figure EC.19). Correct answers were shown after each guess made by the subjects.

To prevent long inactivity during the game, subjects were asked to make all guesses within 30 seconds (in any period of any round). If no guess was made before this time limit, a guess was made automatically, replicating the most recent guess or choosing at random in the first period. Throughout the experiment, all guesses, with no exception, were made by subjects within this time limit.

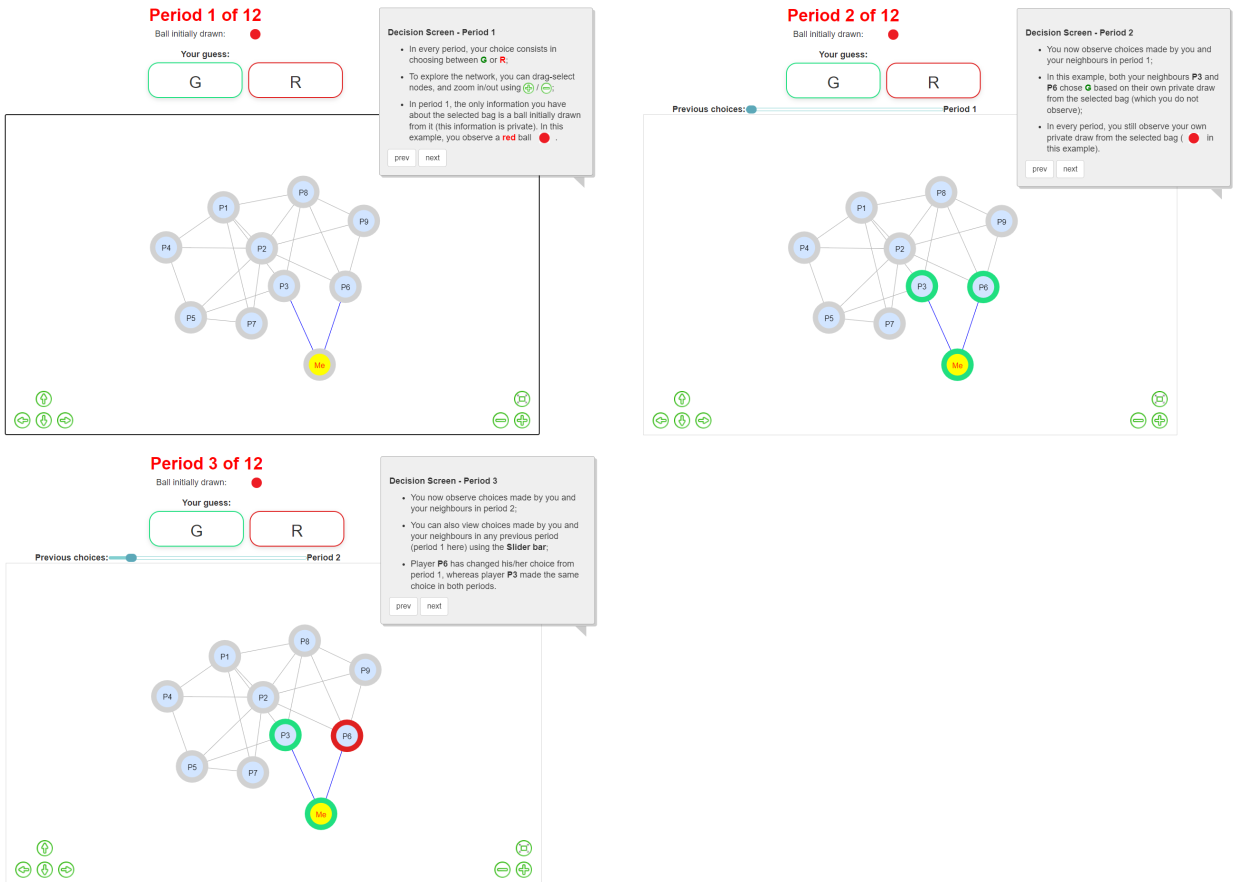


Figure EC.18 **Tutorials from the experiment**

EC.5. Sample Instructions

[The following instructions were used across all treatments.]

Please read the following instructions carefully. **These instructions are the same for all the participants.** The instructions state everything you need to know to participate in the experiment.

In addition to the 5 euro show up fee that you are guaranteed to receive if you complete the experiment, you can earn money based on your choices during the experiment. The other participants will not see how much you earn in the experiment.

There are 6 independent rounds. A round consists of 12 decision periods.

At the very start, before round one, you will be grouped with 39 other participants: there are 40 participants in your group. The composition of the group remains unchanged throughout the 6 rounds.

At the beginning of a round, participants in your group will be randomly assigned to one of 40 positions in a network. A line segment between any two positions represents that they are **connected: Participants in these connected positions observe each other's decision.**

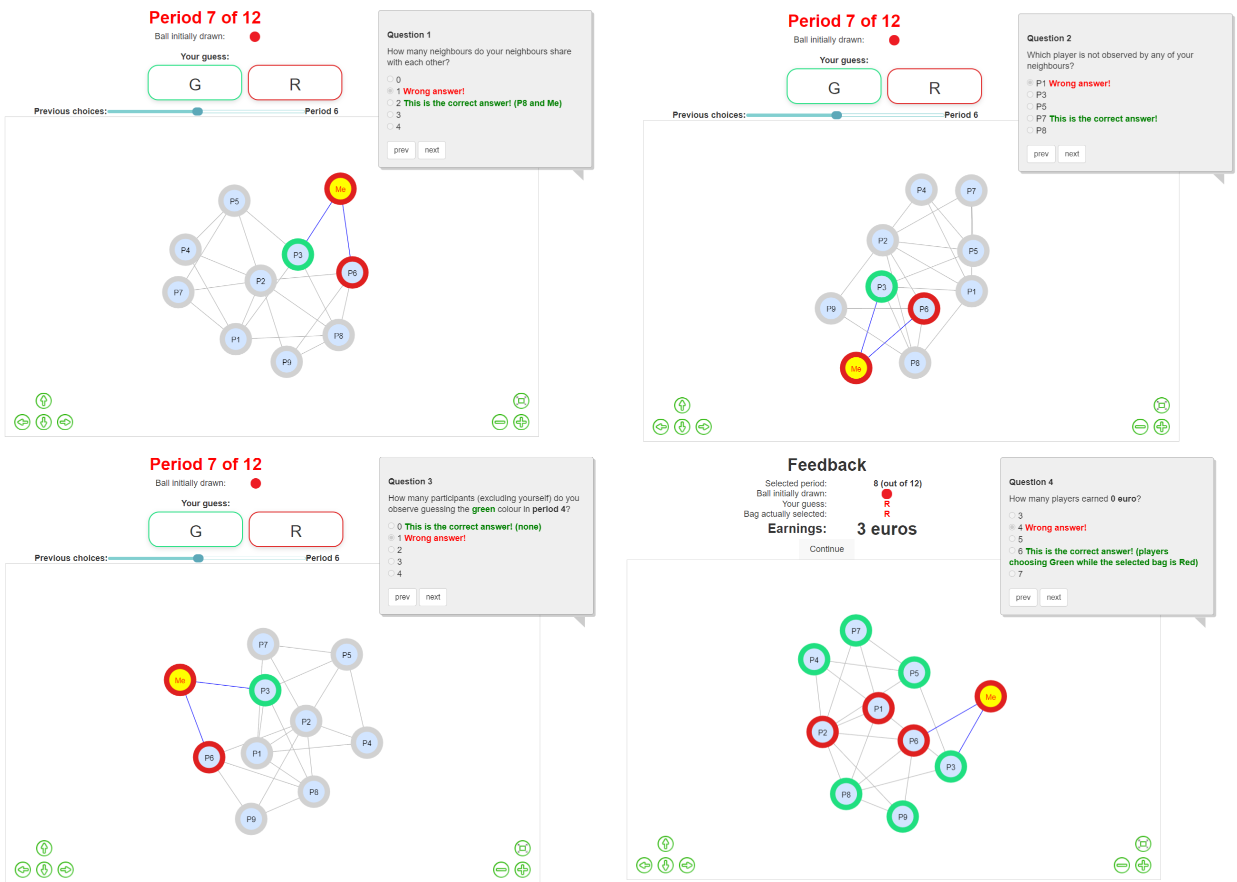


Figure EC.19 Questionnaires from the experiment

Your position in the network is labelled “Me” and other participants are labelled “P1” to “P39”. Your neighbours are participants connected to your position. This assignment remains fixed across the 12 periods of a round.

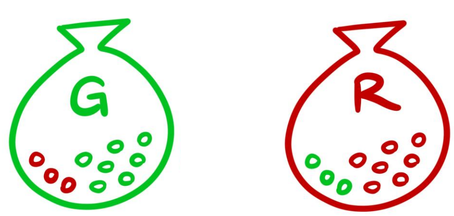
The tutorials in your computer present the network that is used in this experiment. Please go over the tutorials to familiarize yourself with the network.

The assignment of positions in the network depends solely upon chance and is drawn afresh at the start of a round. That is, in each round, every participant is equally likely to be assigned to any position in the network.

A round

At the beginning of a round, **one of these two bags is randomly selected with equal probability.**

Figure EC.20 **Screenshots from the experiment during the game**



The green bag contains 7 green balls and 3 red balls; the red bag contains 3 green balls and 7 red balls. In each of the 12 periods, all participants make a guess of the colour of bag picked.

To help participants make the correct guess, every participant draws one ball from the bag at random. After each draw, the ball will be returned to the bag before making the next private draw.

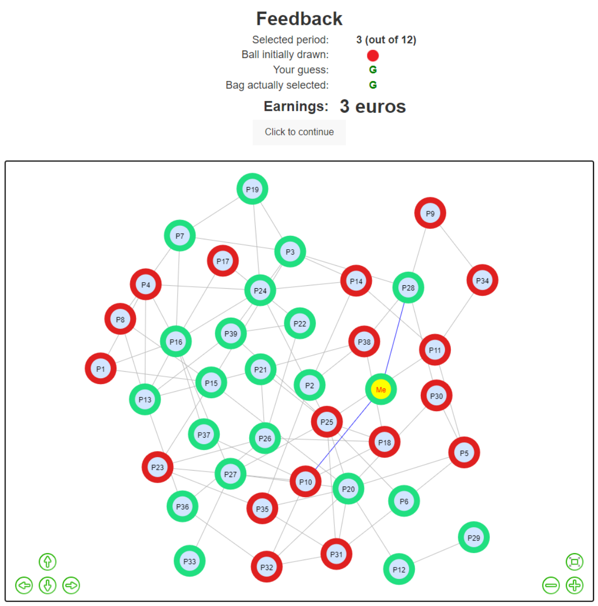
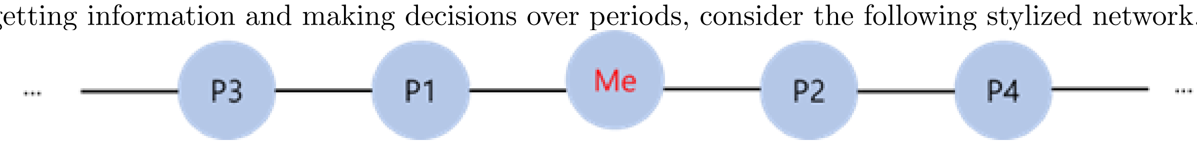


Figure EC.21 **Feedback screen from the experiment during the game**

The colour of the ball drawn for a participant is only observed by that participant and is not shared with anyone else. On top of the computer screen, you will be informed of whether the ball drawn by the computer for you is green or red.

In the first period, after observing his/her private draw, every participant is asked to guess the colour of the bag (**green** or **red**). When ready, you use the mouse to click on the **G(reen)** button or the **R(ed)** button. You are given up to 30 seconds to make your guess.

After all decisions are made, everyone will get to see the choices of their neighbours in the network, i.e., participants to whom s/he is connected in network. In order to explain the process of getting information and making decisions over periods, consider the following stylized network.



After everyone makes a decision in the first period,

- You observe the guesses of P1 and P2
- Your neighbour P1 observes the guesses of you (labelled Me), and P3
- Your neighbour P2 observes the guesses of you and P4
- P3 observes the guesses of P1 and . . .
- P4 observes the guesses of P2 and . . .

As explained in the tutorials, guesses of your neighbours in the previous period are shown through the colours of their node (**green** or **red**).

Next, in period 2, you are asked to guess the colour of the bag, based on your initial private draw and your observation of the guesses made by your neighbours in period 1.

When all participants make their guess in the second period,

- You observe the revised guesses of P1 and P2
- Your neighbour P1 observes the revised guesses of you and P3
- Your neighbour P2 observes the revised guesses of you and P4
- P3 observes the revised guesses of P1 and ...
- P4 observes the revised guesses of P2 and ...

In period 3, you are again asked to guess the colour of the bag. This process is repeated until 12 decision periods are completed.

In every period, you need to make a choice within **30 seconds**. If you do not make any guess within 30 seconds:

- In the first period: the computer will guess randomly for you (green or red with 50% probability).
- In other periods: the computer will carry the guess you made in the previous period.

Note that **you will receive no earnings for guesses made by the computer**.

Recall that the bag chosen in the beginning of a round remains unchanged across the 12 periods of that round. When the first round ends, the computer informs all participants of the bag that was actually chosen. After letting you observe the results of the first round, the second round will start by having the computer randomly assigning network positions to participants in your group and randomly selecting a bag.

This process will be repeated until all 6 independent rounds are completed. At the end of the last round, you will be informed that the experiment has ended.

Earnings and feedback

At the end of each round, one of the 12 periods is randomly picked to determine your earnings for that round. You earn

- 3 Euros if your guess in that period matches the bag picked;
- 0 Euros for an incorrect guess or if you did not make a guess in that period. On the feedback screen, you are provided information on:
- The actual colour of the bag selected;
- Your private draw of a ball from the selected bag;
- The period randomly selected for payment;
- Your guess as well as everyone's guess in the selected period;
- Your earnings.

Your final earnings in the experiment will consist of the sum of earnings across the 6 rounds plus a show-up fee of 5 Euros.

EC.6. Raw data

Figures EC.22 to EC.24 present the evolution of the average guesses of each network treatment (ER, SB, RF), group (1-4), and round (1-6) from the experiment.

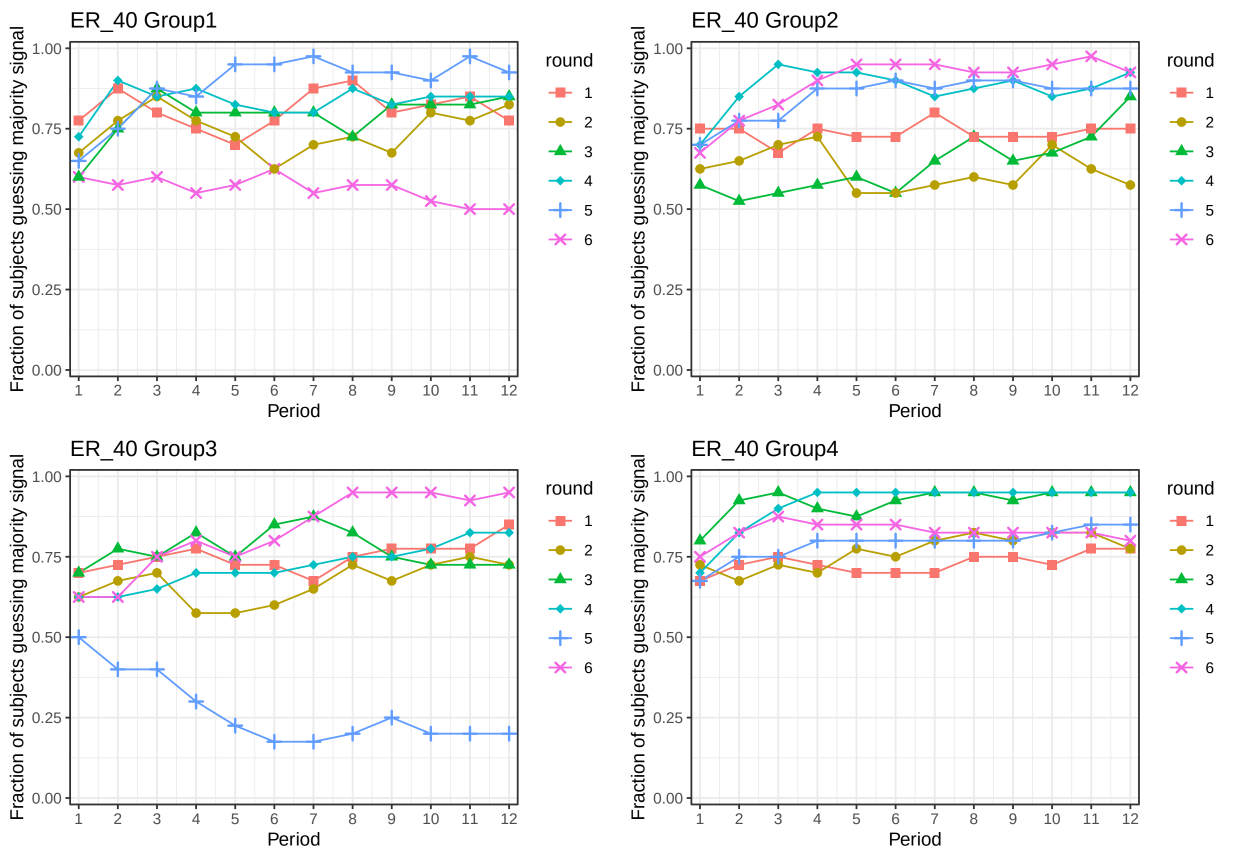


Figure EC.22 Experimental results — Development of guesses n=40

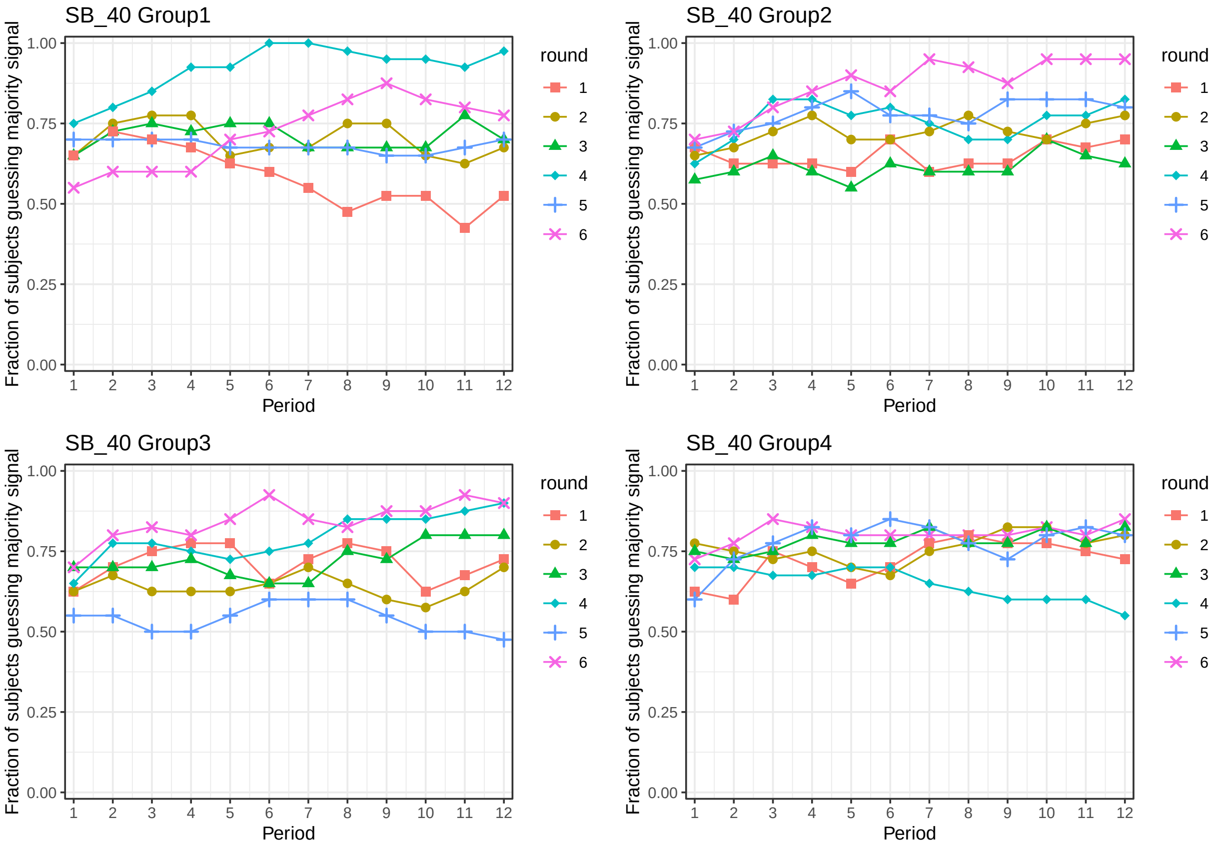


Figure EC.23 Experimental results — Development of guesses n=40

Table EC.6 OLS regression c_t, k=0.2, n=40

	OLS - Correct Consensus	OLS - Incorrect Consensus	OLS - Breakdown
(Intercept)	0.71^{***}	0.04	0.25^{***}
	(0.07)	(0.04)	(0.04)
typeRF	-0.00	0.17^{**}	-0.17^{*}
	(0.15)	(0.08)	(0.09)
typeSB	-0.33^{***}	-0.00	0.33^{***}
	(0.08)	(0.05)	(0.06)
R^2	0.10	0.07	0.20
Adj. R^2	0.08	0.04	0.18
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table EC.7 OLS regression c_t, k=0.3, n=40

	OLS - Correct Consensus	OLS - Incorrect Consensus	OLS - Breakdown
(Intercept)	0.54^{***}	0.04	0.42^{***}
	(0.07)	(0.04)	(0.04)
typeRF	0.08	0.17^{**}	-0.25^{**}
	(0.16)	(0.08)	(0.11)
typeSB	-0.21^{***}	-0.04	0.25^{***}
	(0.07)	(0.04)	(0.04)
R^2	0.06	0.11	0.17
Adj. R^2	0.03	0.08	0.15
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table EC.8 OLS regression c_t, k=0.4, n=40

	OLS - Correct Consensus	OLS - Incorrect Consensus	OLS - Breakdown
(Intercept)	0.46^{***}	0.04	0.50^{***}
	(0.09)	(0.04)	(0.06)
typeRF	0.04	0.12^{*}	-0.17
	(0.18)	(0.07)	(0.12)
typeSB	-0.25^{**}	-0.04	0.29^{***}
	(0.12)	(0.04)	(0.09)
R^2	0.07	0.08	0.14
Adj. R^2	0.04	0.05	0.12
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table EC.9 Logistic regression c_t, k=0.3, n=40

	Logit - Correct Consensus	Logit - Incorrect Consensus	Logit - Breakdown
(Intercept)	0.17	-3.14^{***}	-0.34^{*}
	(0.28)	(0.92)	(0.17)
typeRF	0.34	1.80^{*}	-1.27^{*}
	(0.66)	(1.01)	(0.77)
typeSB	-0.86^{***}	-16.43^{***}	1.03^{***}
	(0.28)	(1.05)	(0.17)
AIC	101.41	38.88	90.78
BIC	108.24	45.71	97.61
Log Likelihood	-47.71	-16.44	-42.39
Deviance	95.41	32.88	84.78
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table EC.10 OLS regression c_t censored, n=40

	OLS - Correct Consensus	OLS - Incorrect Consensus	OLS - Breakdown
(Intercept)	0.36^{***}	0.02	0.62^{***}
	(0.03)	(0.02)	(0.01)
typeRF	0.07	0.08^{**}	-0.15^{*}
	(0.10)	(0.04)	(0.08)
typeSB	-0.14^{***}	-0.02	0.16^{***}
	(0.04)	(0.02)	(0.02)
R^2	0.08	0.09	0.19
Adj. R^2	0.06	0.06	0.17
Num. obs.	72	72	72

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table EC.11 OLS regression of welfare w_t on network treatment

	OLS - Welfare
(Intercept)	0.30^{***}
	(0.07)
typeRF	-0.12
	(0.17)
typeSB	-0.18^{**}
	(0.08)
R^2	0.02
Adj. R^2	0.01
Num. obs.	432

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

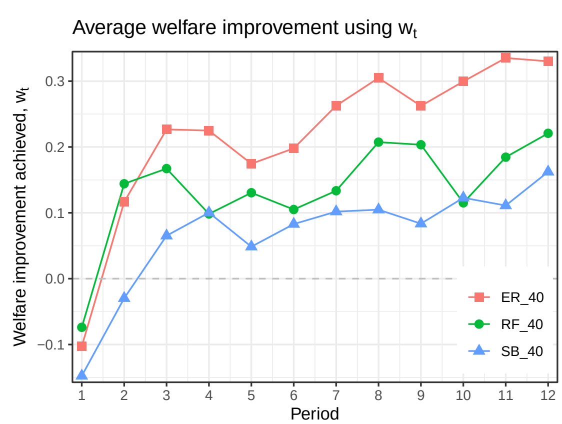


Figure EC.10 Evolution of welfare improvement w_t in the experiment. In the experiment, in the last period, ER achieved 30% of the possible welfare improvement, 18% for SB and 22% for RF.

Table EC.12 Fraction of guesses imitate leader against DeGroot prediction

	Correctly follow leader			
	OLS (Bayesian predicts 1)	Logit	OLS	OLS
(Intercept)	0.10^{***}	-2.20^{***}	0.18^{***}	0.18^{***}
	(0.02)	(0.18)	(0.03)	(0.02)
RF_40	-0.06^{***}	-0.91^{**}		
	(0.02)	(0.39)		
SB_40	0.04^{**}	0.41^{*}		
	(0.02)	(0.21)		
period			-0.01^{***}	
			(0.00)	
round				-0.02^{***}
				(0.01)
R^2	0.01		0.01	0.01
Adj. R^2	0.01		0.00	0.01
Num. obs.	1870	1870	1870	1870
AIC		1388.23		
BIC		1404.83		
Log Likelihood		−691.12		

$^{***}p < 0.01$; $^{**}p < 0.05$;$^{*}p < 0.1$

Table EC.13 Fraction of guesses following signal against DeGroot prediction

	Always follow signal			
	OLS (Stubbornness predicts 1)	Logit	OLS	OLS
(Intercept)	0.25^{***}	-1.07^{***}	0.35^{***}	0.40^{***}
	(0.01)	(0.07)	(0.02)	(0.02)
RF_40	0.04	0.21		
	(0.04)	(0.21)		
SB_40	0.04^{*}	0.22^{*}		
	(0.02)	(0.12)		
period			-0.01^{***}	
			(0.00)	
round				-0.03^{***}
				(0.01)
R^2	0.00		0.00	0.01
Adj. R^2	0.00		0.00	0.01
Num. obs.	9366	9366	9366	9366
AIC		11185.38		
BIC		11206.81		
Log Likelihood		−5589.69		

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Nb of checks	ER	RF	SB
0	55%	58%	44%
≥1	45%	42%	56%

(a) Frequency of subjects choosing to view the outcome of earlier periods (t-2, t-3,...) at least once during the experiment.

Nb of checks	ER	RF	SB
0	97%	96%	95%
≥1	3%	4%	5%

(b) Frequency of guesses for which subjects chose to view the outcome of earlier periods (t-2, t-3,...).

Table EC.14 Activity of viewing the outcome of earlier periods (t-2, t-3,...) for each of the network treatments (by default, the outcome of the previous period t-1 is shown to subjects).

	ER		RF		SB	
	% DeGroot	N	% DeGroot	N	% DeGroot	N
View earlier periods	87%	251	81%	329	84%	395
Only view t-1	89%	8817	86%	9249	86%	8657
Mann-Whitney U	1088451		1459631		1680777	
Mann-Whitney p	0.42		0.04		0.35	

Table EC.15 Frequency of guesses matching the DeGroot decision rule conditional on whether the corresponding subject viewed the outcome of earlier periods (t-2, t-3,...), across each network (data points corresponding to the first 2 periods of each round, and data points where the DeGroot rule predicts indifference are excluded from this analysis).

Table EC.16 Quartile and Mean of c_t under DeGroot simulation.

type	1st quartile	2nd quartile	3rd quartile	mean
RF	1	1	1	0.792
ER	0.95	1	1	0.953
RGG	0.825	0.925	1	0.882
SB	0.75	0.875	1	0.864

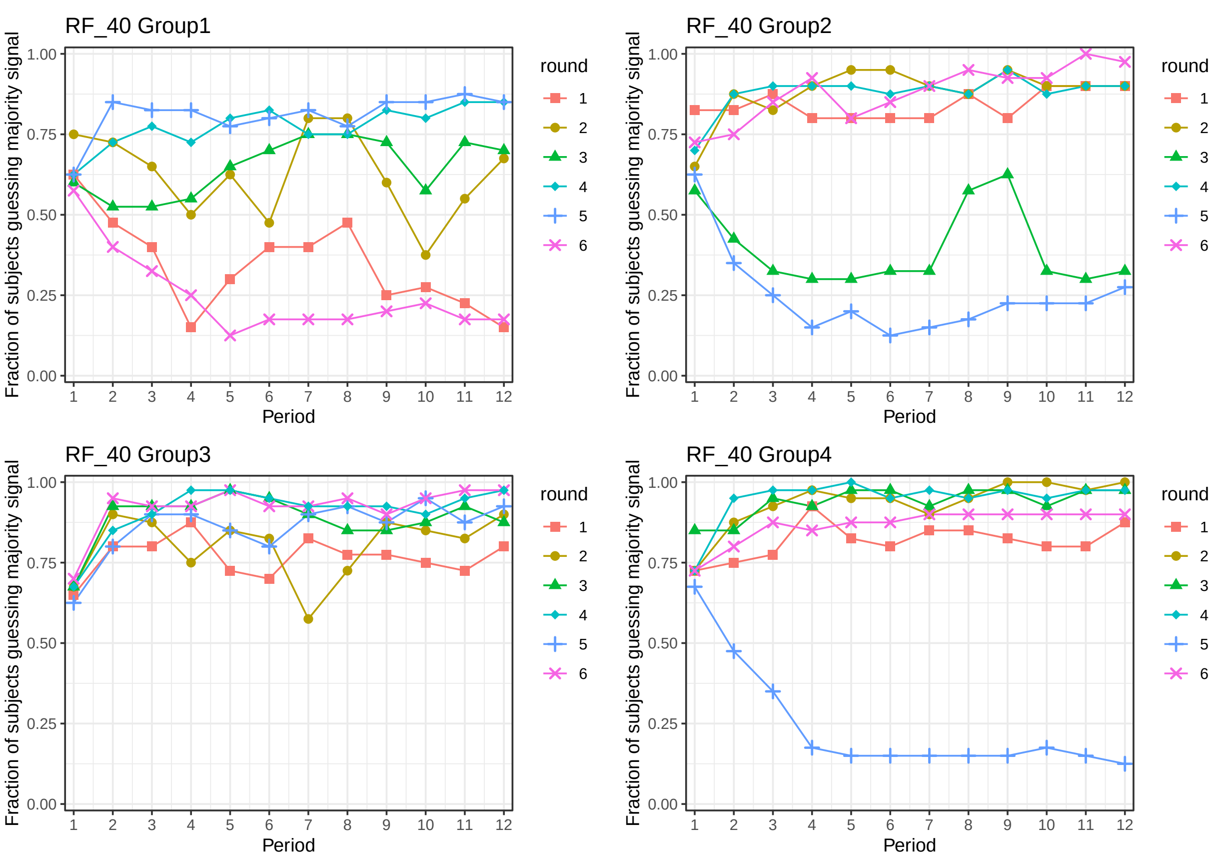


Figure EC.24 Experimental results — Development of guesses n=40