ONLINE Appendix for "Learning through Imitation: an Experiment"

Marina Agranov* Gabriel Lopez-Moctezuma†
Philipp Strack‡ Omer Tamuz§

A Structure of the ALL treatment and payments

Each participant earns $10 for completing the session (participation fee). In addition, subjects earn money for other parts as described below.

- Part I: Main game
  - 10 games with 20 rounds in each game
  - random re-matching into groups of 8 subjects at the beginning of a game
  - random round from a random game determines participants' payments for this part: correct guess pays $20, wrong guess pays $5
- Part II: Beliefs
  - one questions is randomly selected for payment[1], $5 for a correct answer
  - questions appear in a random order for each subject
  - in the first session of each treatment, this part was not present because we were collecting the data used in the next sessions for subjects' payments.
- Part III: Risk Attitudes

*Caltech. Email: magranov@hss.caltech.edu.
†Caltech. Email: glmoctezuma@caltech.edu.
‡Yale University. Email: philipp.strack@yale.edu.
§Caltech. Email: tamuz@caltech.edu.

[1]For the NO INFO, ACTIONS, and ALL treatment we had 3 beliefs questions corresponding to these three information structures. In the SIGNALS treatment, we had 4 beliefs questions each corresponding to one of the all four information structures.

  - one of the two investment tasks is randomly selected for payment
  - points earned are converted into dollars using the rate 1 point = 1 cent

- Part IV: IQ and Overconfidence
  - six matrices for IQ measure, each correctly solved matrix earns 50 cents
  - overconfidence is measured using two related dimensions: over-estimation and over-placement
  - correct prediction for over-estimation task earns 50 cents
  - in the first session of each treatment (those without beliefs questions) over-placement question is not incentivized, while in the remaining sessions it is and correct prediction is rewarded by 50 cents[2]

- Part V: Open-ended Questionnaire for Strategies

B Instructions for ALL treatment

Welcome. You are about to participate in an experiment on decision-making. You will be paid for your participation in cash privately at the end of the session. Please turn off all electronic devices, especially phones. During the experiment you are not allowed to open or use any other applications on these laboratory computers, except for the interface of the experiment.

The experiment consists of four parts: Part I of the experiment is the main and the longest part. The other parts (Part II, III and IV) are short. You will receive the instructions for each part of the experiment before that part begins. You have already earned $7 for coming to the lab. In addition, you can earn money in each part of the experiment. The instructions for each part of the experiment will be very precise about that.

Part I

Part I of the experiment consists of 10 games. Each game consists of 20 rounds. Before the beginning of each game, you will be randomly assigned to a group of 8 players. You will play all 20 rounds of the game with the same group of people. At the end of each game, you will again be randomly assigned to a new group of 8 players and will play 20 rounds with them, and so on.

[2]The first session is used to collect data for comparing actual rank of students to that reported for the follow-up sessions.

At the beginning of each game, the computer randomly selects one of the two Urns for your group for this game (independent of the urns selected in previous games):

- RED URN contains 6 RED balls and 4 GREEN balls
- GREEN URN contains 4 RED balls and 6 GREEN balls

That is, there is 50% chance that the RED URN is selected, in which case the urn has 6 RED balls and 4 GREEN balls and 50% chance that the GREEN URN is selected, in which case the URN has 6 GREEN balls and 4 RED balls. We will refer to the selected urn as the URN.

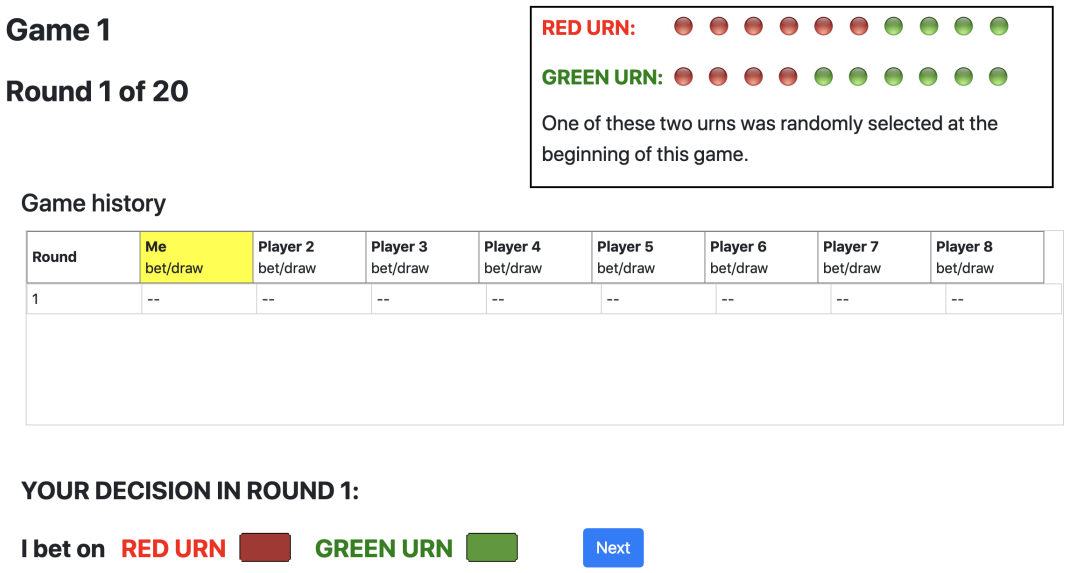
The composition of the URN is determined once at the beginning of each game (before round 1) and stays the same throughout the game (in all 20 rounds). All players that were assigned to the same group share the same composition of the URN. At the beginning of a new game, after players are assigned to new groups but before round 1 begins, the computer again determines the composition of the URN for each group separately using the rule described above. Thus, the composition of the URN in your group stays the same in all 20 rounds of the same game but is not related to its composition in later games.

In every round of a game, your task is to guess (bet on) the URN selected for your group at the beginning of this game. If you choose to bet on RED URN, it means that you are betting that the RED URN was selected for your group at the beginning of this game. If you choose to bet on GREEN URN, it means that you are betting that the GREEN URN was selected for your group at the beginning of this game. At the end of each round, one ball will be randomly drawn from the selected URN and its color revealed to you. The same thing happens with all other players in your group: each player bets on the URN selected for your group and then observes one randomly drawn ball from the selected URN. It is important that different players observe different draws, but that all drawn are made from the same URN.

Your payment in Part I. To determine your payment in Part I, at the end of the experiment, the computer will select one game from the 10 games played. Each game is equally likely to be chosen for payment. Then the computer will select one of the 20 rounds in the selected game. If you guessed correctly the URN in the selected round of the selected game, then you will receive $20. If your guess (bet) was wrong then you will receive $5.

End-of-Round Information. At the end of each round, you will be reminded of the URN you betted on and you will observe the color of the randomly drawn ball from the URN. Moreover, at the end of each round you will also observe the bets that

Figure B.1: Screenshot 1 from All treatment



other players in your group made in this round regarding the selected URN and the balls that were drawn from the selected URN for other players. This information will be summarized in the table on your screen. This table will keep track of all decision that you have made in this game (highlighted in yellow), as well as the decisions made by other players in your group.

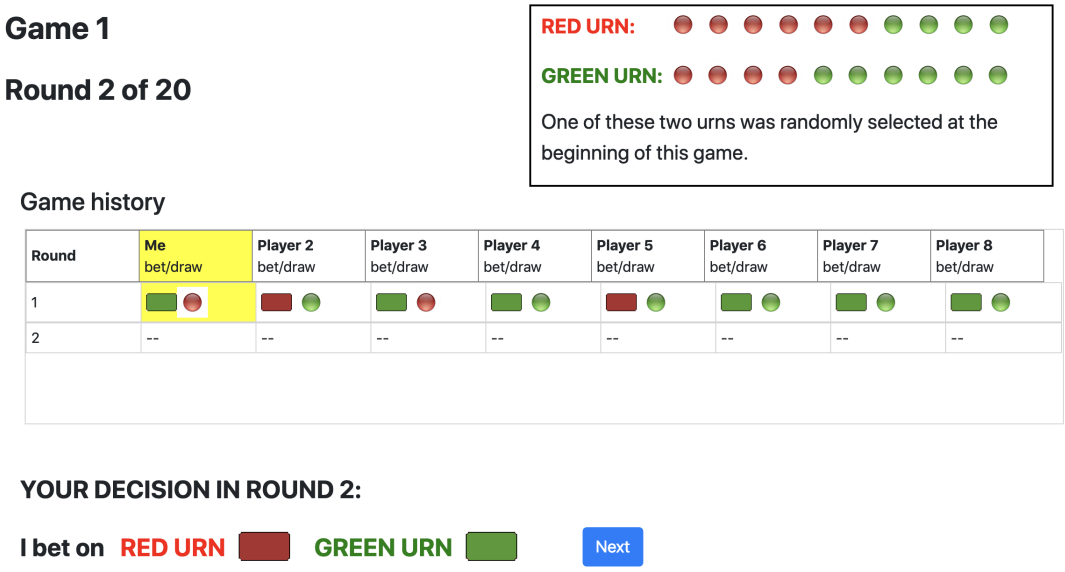
Note that the computer draws one ball for each member of your group in each round from the URN with replacement. That is, every ball that is drawn from the URN is placed back in the URN before the next draw. Therefore, if the URN contains, say, 6 RED balls and 4 GREEN balls then for each player in your group there is exactly 60% chance that the drawn ball is RED and 40% chance that it is GREEN.

Are there any questions?

B.1 Screenshots for ALL treatment

This screenshot presented in Figure B.1 shows round 1 of a game. On the top of the screen, subjects are reminded about the compositions of two urns. The game history table keeps track of all what has transpired in the current game, and the bottom of the screen is where subjects make their bets about which urn was selected for their group for this game.

Figure B.2: Screenshot 2 from All treatment



The game history table starts filling up at the beginning of round 2 (as shown in Figure B.2), and it displays own bet and own draw as well as bets and draws of all group members (as seen on the second screenshot). Bets are always displayed as rectangles, i.e., lottery tickets, and signals are displayed as circles, i.e., balls drawn from the selected urn.

B.2 Strategy questions in ALL treatment

Think about the main game in the experiment (betting on the urn selected for your group).

- What strategy did you use in the game (if any)? Please elaborate.
- Did you look at the bets made by other players in your group? Did you find them useful/not useful? Please elaborate.
- Did you look at the balls drawn for other players in your group? Did you find them useful/not useful? Please elaborate.
- Was anything unclear about the game?
- What is your gender?
- What is your major?

B.3 Beliefs questions in ALL treatment

Below we present the three beliefs questions that subjects in the ALL treatment were asked to answer. The formulation for the beliefs questions in the other treatments is adapted to the information structure subjects experienced in the treatment, but follow the same idea. The questions were presented in the random order across subjects.

Question 1. In the past session, groups of 8 subjects played the same game that you just played. Namely, at the beginning of each game, the computer randomly selected one of the two urns: with probability 50% the RED URN was selected, which contained 6 RED balls and 4 GREEN balls, and with probability 50% the GREEN URN was selected, which contained 4 RED balls and 6 GREEN balls. Each game lasted for 20 rounds. In each round group members submitted their guesses (bets) about the URN selected for their group for this game. After making their choices, each group member observed the colors of the 8-randomly drawn balls from the selected urn (with replacement) and also guesses made by members of their group regarding the selected urn for this game.

Please answer the following question: What fraction of last round bets (bets submitted by subjects in round 20) were correct? The correct guess means that a subject guessed correctly which URN was selected for her group for this game.

[[Radio buttons with ten options: 0%-9%, 10%-19%, ..., 90% - 100%]]

If this question is selected for payment, you will get paid $5 if you chose correctly.

Question 2. In the past session, groups of 8 subjects played game that was slightly different from the game you just played. Just like in your game, at the beginning of each game, the computer randomly selected one of the two urns: with probability 50% the RED URN was selected, which contained 6 RED balls and 4 GREEN balls, and with probability 50% the GREEN URN was selected, which contained 4 RED balls and 6 GREEN balls. Each game lasted for 20 rounds. In each round group members submitted their guesses (bets) about the URN selected for their group for this game. After making their choices, each group member observed the colors of the one randomly drawn ball from the selected urn and the guesses made by members of their group regarding the selected urn.

However, contrary to the game you just played, participants in these previous sessions did not observe the colors of the balls that were randomly drawn for other participants in their group. These past participants only observed the guesses that their group members made in each round and one randomly drawn ball from the selected urn.

Please answer the following question: What fraction of last round bets (bets

submitted by subjects in round 20) were correct? The correct guess means that a subject guessed correctly which URN was selected for her group for this game.

[[Radio buttons with ten options: 0%-9%, 10%-19%, ..., 90% - 100%]]

If this question is selected for payment, you will get paid $5 if you chose correctly.

Question 3. In the past session, groups of n (parameter in the program) subjects played game that was slightly different from the game you just played. Just like in your game, at the beginning of each game, the computer randomly selected one of the two urns: with probability 50% the RED URN was selected, which contained 6 RED balls and 4 GREEN balls, and with probability 50% the GREEN URN was selected, which contained 4 RED balls and 6 GREEN balls. Each game lasted for 20 rounds. In each round group members submitted their guesses (bets) about the URN selected for their group for this game. After making their choices, each group member observed the colors of the one randomly drawn ball from the selected urn.

However, contrary to the game you just played, participants in this previous session were not provided with any additional information about the guesses made by other members of their group or the colors of the randomly drawn balls from the selected urn for other participants. The only information these past participants had was the color of one randomly drawn ball from the selected urn.

Please answer the following question: What fraction of last round bets (bets submitted by subjects in round 20) were correct? The correct guess means that a subject guessed correctly which URN was selected for her group for this game.

[[Radio buttons with ten options: 0%-9%, 10%-19%, ..., 90% - 100%]]

If this question is selected for payment, you will get paid $5 if you chose correctly.

B.4 Other Control Tasks

Risk Attitudes. Risk attitudes were measured using two investment tasks, in each of which subjects were endowed with 200 points (worth a total of $2), any portion of which they could choose to invest in a risky project. In the first investment task, the risky project was successful 50% of the time and had a return of 2.5 points for each point invested in it, while in the second investment task the risky project was successful 40% of the times and returned 3 points for each point invested in it. Points not invested in the risky project had a return of 1 to 1 point. One of these two tasks was randomly selected for payment. This is one of the standard methods used

in the experimental literature to elicit subjects' attitudes towards risk (see Gneezy and Potters (1997) and Charness, Gneezy, and Imas (2013)). Administering this task twice with two sets of parameters allows to reduce measurement error (see ORIV technique developed by Gillen, Snowberg, and Yariv (2018)).

IQ and Overconfidence. Subjects were asked to solve six matrices from the ICAR database (see ICAR, Condon and Revelle (2014)). Subjects earned 50 cents for each correctly solved matrix. So, the IQ of a subject is measured by the number of correctly solved matrices with the smallest number being 0 and the largest number being six. After solving these matrices, we asked subjects two questions:

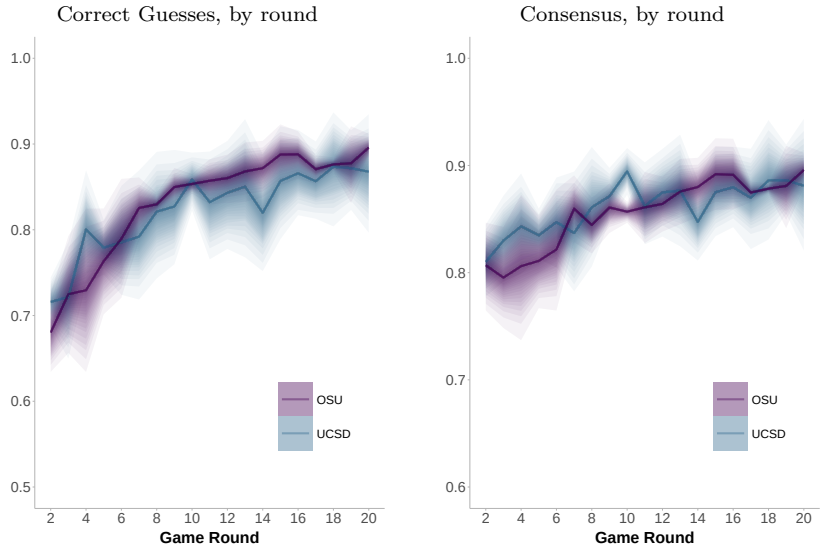
1. How many of the six puzzles do you think you correctly answered? You will receive 50 cents if you answer correctly.
2. Now, think about 100 UCSD students. Where do you think you rank in terms of how many correct rotation cubes puzzles you got? For example, if you think you got the most correct, you should answer 1, while if you think you got the least correct, you should answer 100.

The last two questions are used to measure overconfidence of subjects. We chose to measure over-estimation and over-placement of subjects. Specifically, overestimation is the difference between how many ICAR questions a subject thinks she solved correctly minus how many she actually solved correctly, while the over-placement is the reported rank minus actual rank in a sample of the 100 randomly selected students. These measures are quite standard in the literature (see Chapman et al. 2019b).

C Physical *vs* Virtual Lab

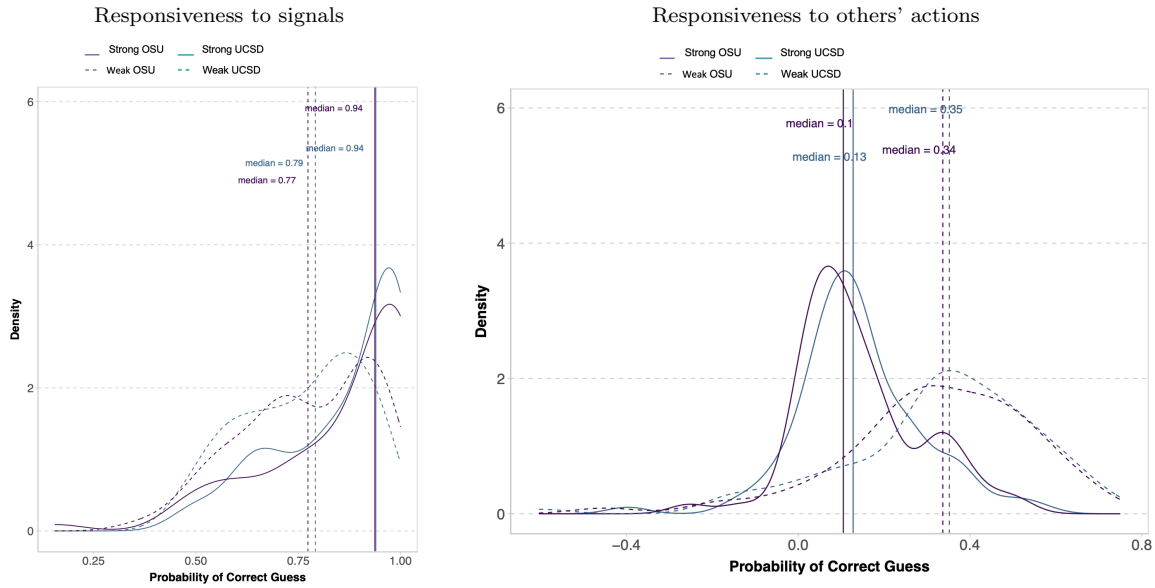
We compare outcomes and behavior observed in the ALL treatment across sessions conducted in the physical lab at UCSD and in the virtual lab at OSU. For each location we have four sessions with 64 subjects all together at UCSD and 88 subjects at OSU. In both locations, we used the standard subjects' pool of undergraduate students.

Figure C.1: Aggregate statistics in the ALL treatment



Notes: Panel (a) presents the average frequency of correct guesses in the ALL treatment in each round, averaged across games. Panel (b) depicts the evolution of consensus in each round, i.e., the relative size of the majority, averaged across games.

Figure C.2: Individual responsiveness to signals and others' actions in ALL treatment



Notes: Panel (a) presents Kernel distributions of participants' responsiveness to signals separately for weak and strong signals. The vertical lines depict median responsiveness for each group. Responsiveness to signals is given by the probability that a participant's action matches signal majority, when actions of others group members in the previous round are split equally between green and red. Panel (b) presents Kernel distributions of participants' responsiveness to others' actions separately for weak and strong signals. The vertical lines depict median responsiveness for each group. Responsiveness to others' actions is measured by the change in the probability of choosing the action of the majority of signals when all versus none of the other group members choose the majority of signals in the last round.

Figure C.1 presents two main outcomes of interest in each location: fraction of correct guesses in Panel (a) and consensus rates in Panel (b). The evolution and the levels of both outcomes are extremely similar in the two locations, which is confirmed by the statistical analysis. Regression analysis detects no significant differences between two locations with $p > 0.10$ in all comparisons. Figure C.2 complements aggregate results by depicting individual responsiveness to signals and to others' actions in the two locations. We find no significant differences between the distribution of individual behavior of subjects in the two locations, as measured by these two statistics.

We conclude by noting that we detect no significant differences in neither aggregate results nor in the individual level results between sessions conducted in person at UCSD and sessions conducted online at OSU, which is why the analysis in the main text pools together these sessions.

D Learning Across Games

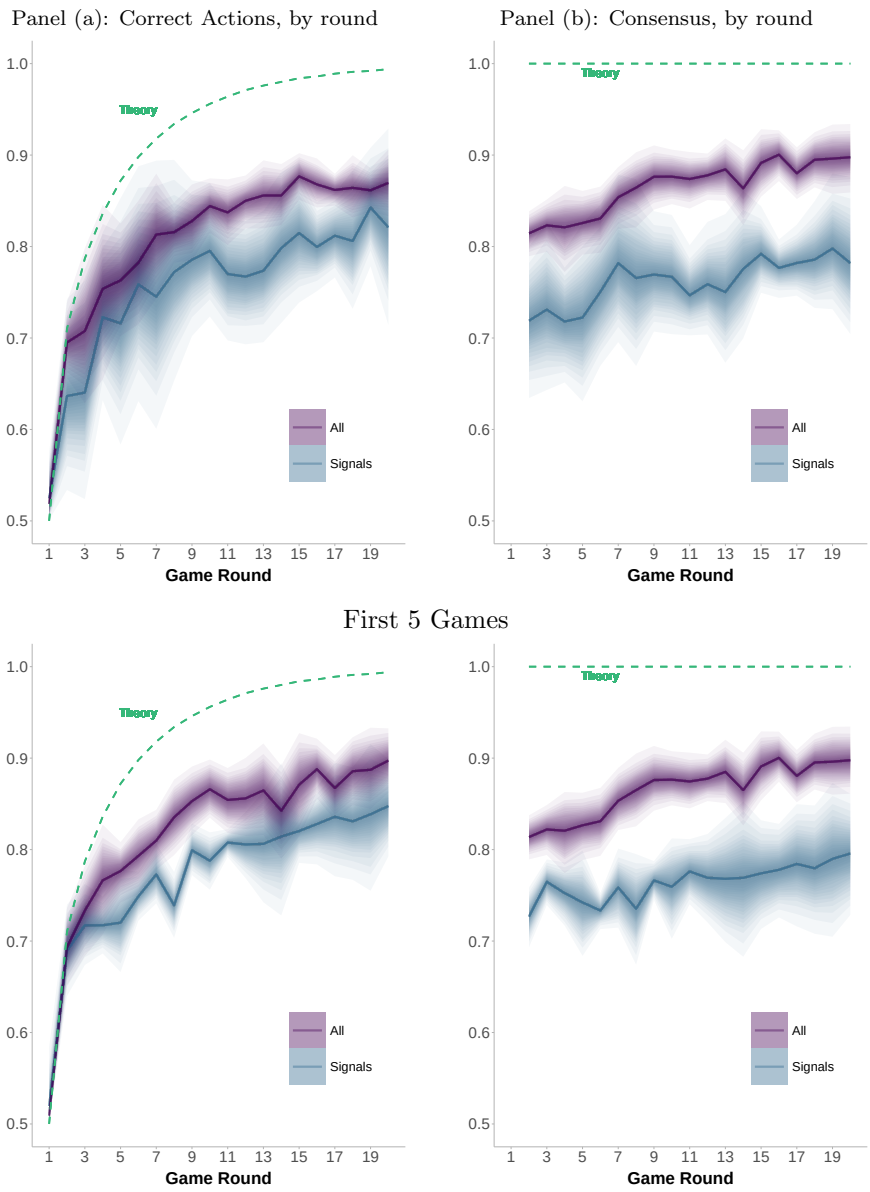


Figure D.1: ALL and SIGNALS treatments (early versus late games)

Notes: Panel (a) presents the average frequency of correct actions in each treatment in each round, averaged across the first 5 games (top figure) and the last 5 games (bottom figure). Panel (b) depicts the evolution of consensus in each round, i.e., the relative size of the majority, averaged across the first 5 games (top figure) and the last 5 games (bottom figure). For Panel (b) we exclude cases with equal number of green and red signals. Shaded regions represent confidence intervals from 50% (darkest) to 95% (faintest) probability levels. Confidence intervals are constructed with a variance-covariance matrix clustered by session.

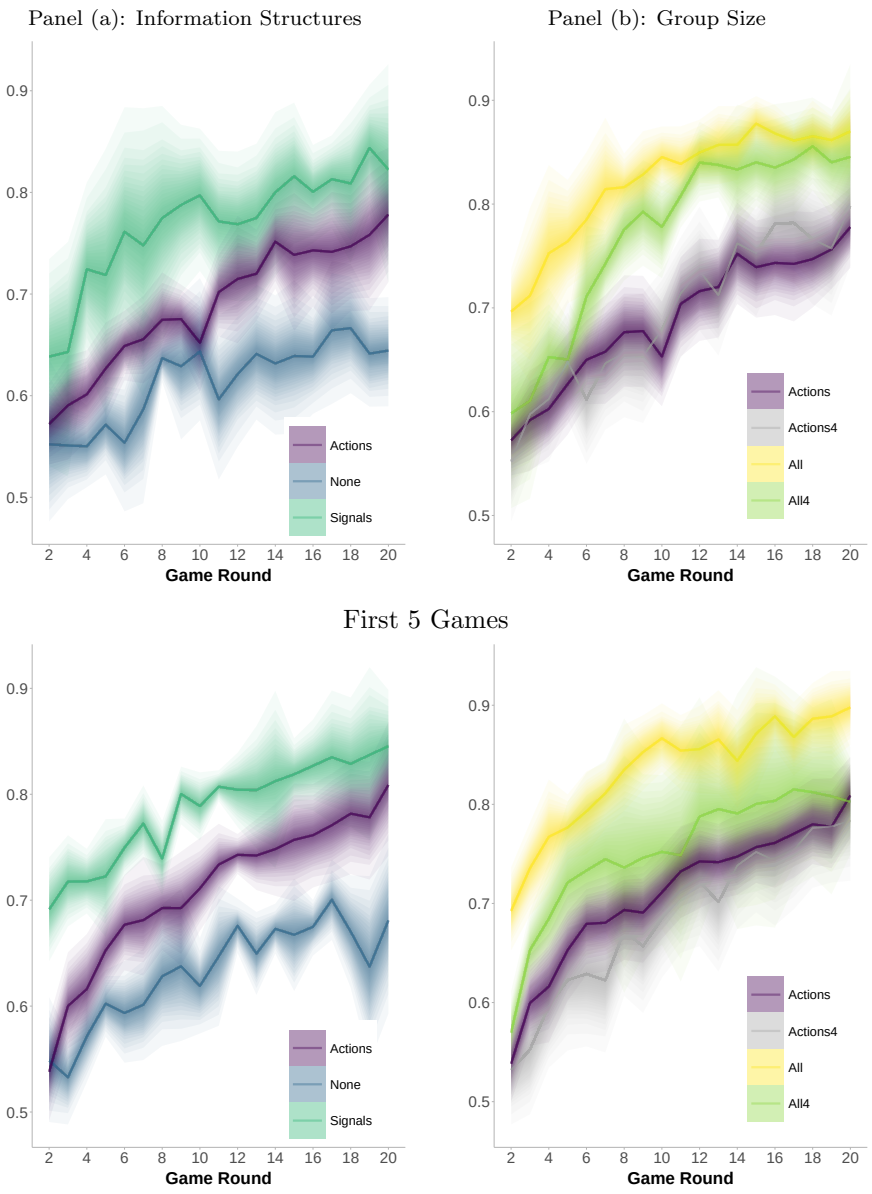


Figure D.2: Frequency of correct actions (early versus late games), by information structure and group size

Notes: Both panels present the average frequency of correct actions in each treatment per each round, averaged across the first 5 games (top figure) and the last 5 games (bottom figure). Shaded regions represent confidence intervals from 50% (darkest) to 95% (faintest) probability levels. Confidence intervals are constructed with a variance-covariance matrix clustered by session.

E Learning From Others' Actions

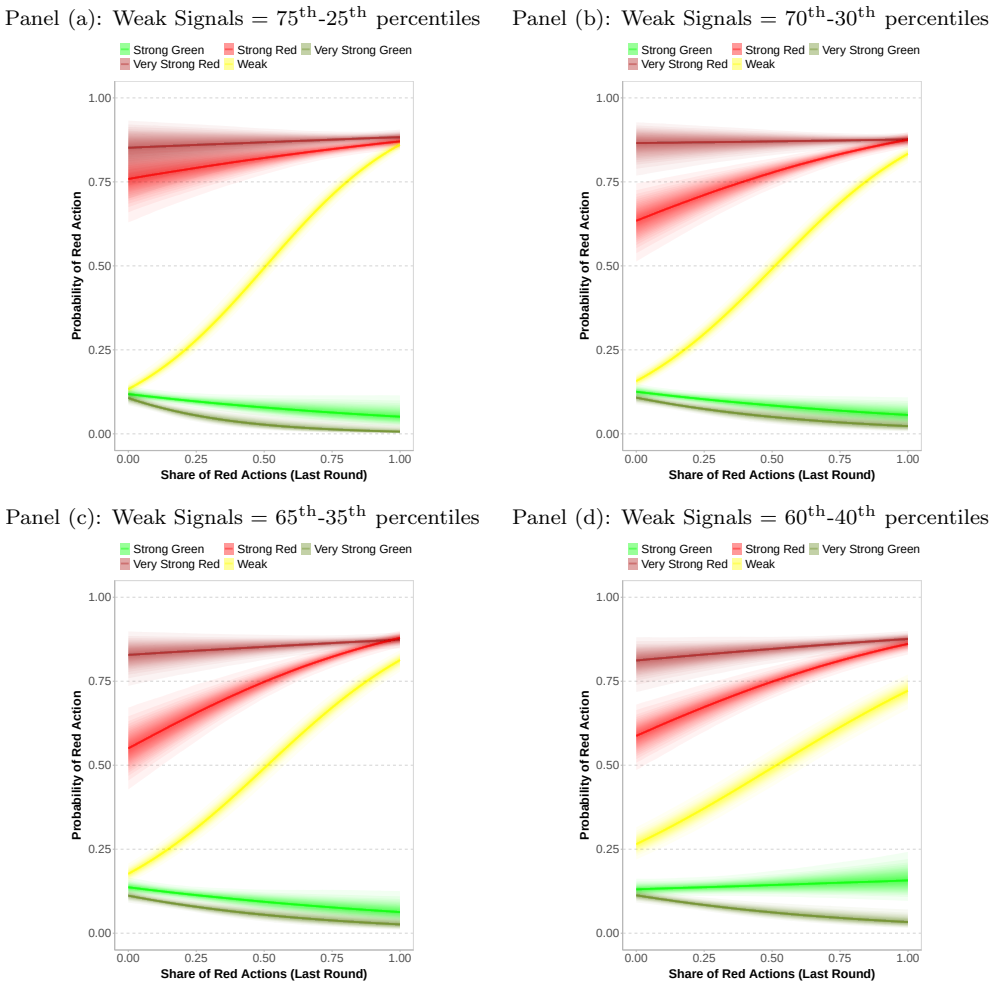


Figure E.1: Different Cutoffs for Signal Strength in ALL treatment

Notes: Each panel depicts the probability of guessing red as a function of the share of red actions of other group members in the previous round. The estimates are obtained from a Bayesian logistic regression of subjects' actions on the share of others' actions in the previous round conditional on signal strength. For panel (a) we use the classification in the main text: *Very Strong Green* with $percentile = (0, 0.1]$, *Strong Green* with $percentile = (0.1, 0.25]$, *Weak* with $percentile = (0.25, 0.75]$, *Strong Red* with $percentile = (0.75, 0.9]$ and *Very Strong Red* with $percentile = (0.9, 1)$. For the cutoffs of panels (b), (c) and (d) we reduce the *Weak* signals category and increase the "strong" categories proportionally: *Very Strong Green* with $percentile = (0, 0.1 + x]$, *Strong Green* with $percentile = (0.1 + x, 0.25 + x]$, *Weak* with $percentile = (0.25 + x, 0.75 - x]$, *Strong Red* with $percentile = (0.75 - x, 0.9 - x]$ and *Very Strong Red* with $percentile = (0.9 - x, 1)$, where $x \in \{0.05, 0.1, 0.15\}$.

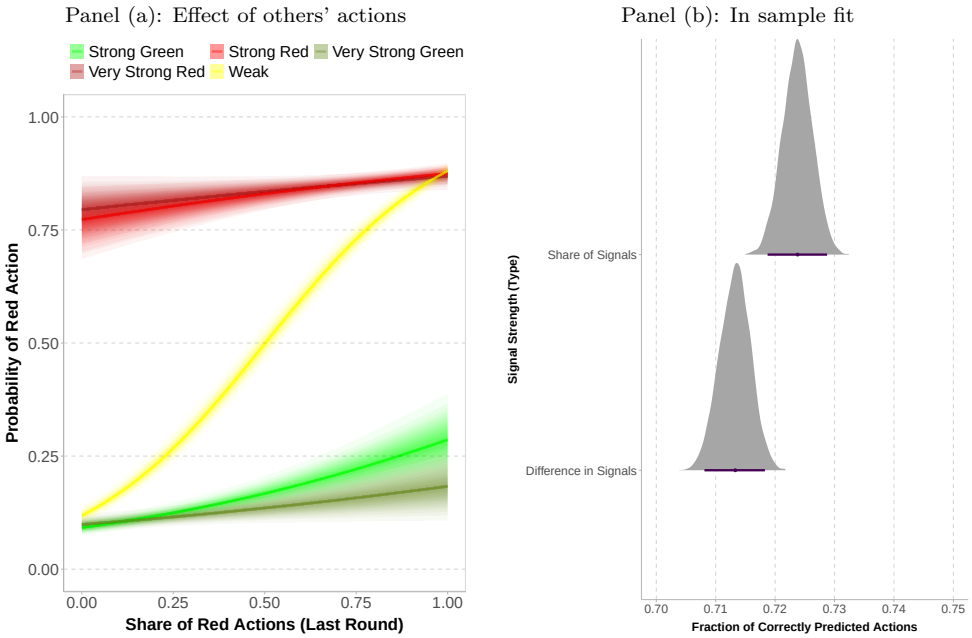


Figure E.2: Alternative Signal Strength Measures (Difference in Signals *vs* Share of Signals)

Notes: Panel (a) depicts the probability of choosing red as a function of the share of red actions of other group members conditional on signal strength as constructed from the share of red signals. Shaded regions represent 95% credible intervals from 50% (darkest) to 95% (faintest) probability levels. Panel (b) depicts the posterior distributions, along with 95% credible intervals, of the fraction of correctly predicted guesses using either the difference in red and green signals or the share of red signals to construct signal strength. For both measures of signal strength, we estimate a Bayesian logistic regression of subjects' actions on the share of others' actions in the previous round conditional on signal strength and session random effects.

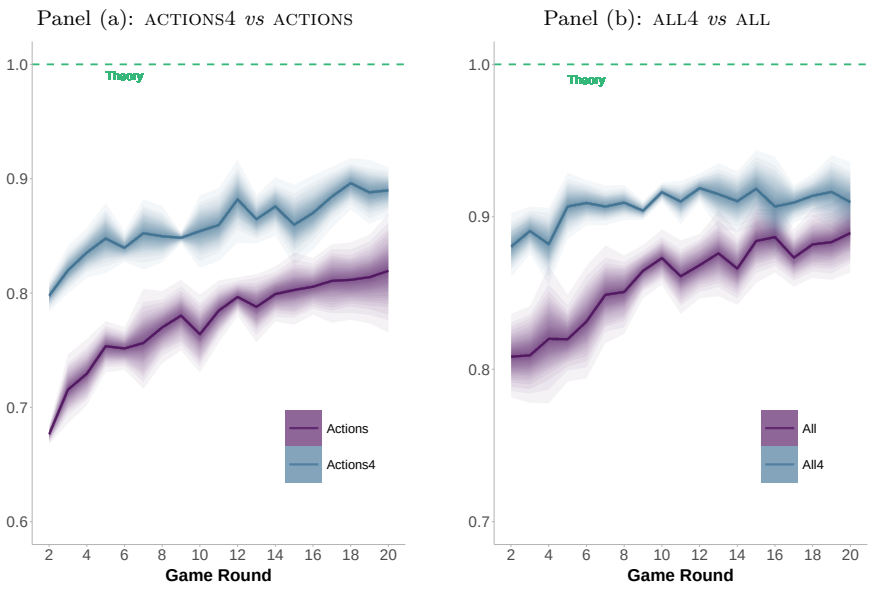
F Additional Aggregate Results

Table F.1: Treatment Effects by Group Size

	Dependent variable:			
	Consensus Rate			
	(1)	(2)	(3)	(4)
ACTIONS (Baseline)	0.686***	0.681***		
	(0.011)	(0.010)		
ALL (Baseline)			0.820***	0.812***
			(0.010)	(0.011)
ACTIONS4 (Effect)	0.090***	0.101***		
	(0.010)	(0.009)		
ACTIONS4 (Effect) × Late Rounds		−0.020*		
		(0.010)		
ALL4(Effect)			0.049***	0.064***
			(0.009)	(0.011)
ALL4 (Effect) × Late Rounds				−0.028**
				(0.014)
Game Round Fixed Effects	Yes	Yes	Yes	Yes
Observations	5,550	5,550	6,841	6,841
Adjusted R^2	0.160	0.161	0.058	0.061

Notes: *p<0.1; **p<0.05; ***p<0.01. Clustered standard errors by session in parentheses. Late rounds are 11-20.

Figure F.1: Consensus Rates, by Group Size



Notes: Both panels present the consensus rates in each treatment per each round, averaged across games within a session. Shaded regions represent confidence intervals from 50% (darkest) to 95% (faintest) probability levels. Confidence intervals are constructed with a variance-covariance matrix clustered by session.

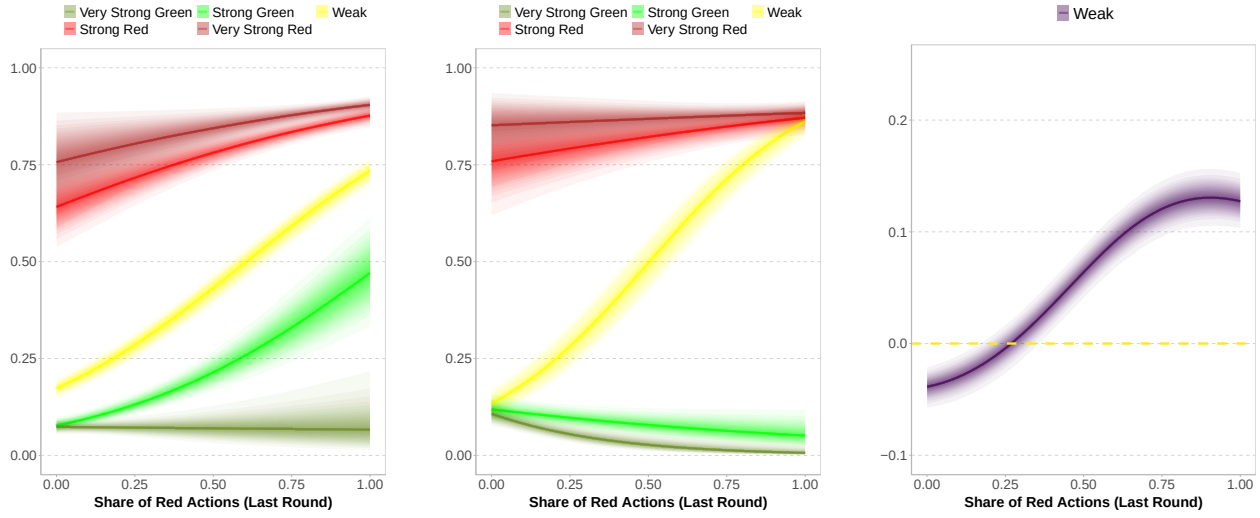
Table F.2: Probability of Red Bet by Signal Strength (ALL versus SIGNALS)

	Dependent variable:			
	Red Bet			
	(1)	(2)	(3)	(4)
Constant	0.088***	0.199***	0.174***	−0.078
	(0.011)	(0.023)	(0.056)	(0.060)
SIGNALS	0.032	0.032	0.046*	0.300***
	(0.032)	(0.032)	(0.026)	(0.026)
Strong Green	0.020**	0.015*	0.032**	0.031*
	(0.009)	(0.009)	(0.016)	(0.016)
Weak	0.369***	0.315***	0.321***	0.306***
	(0.019)	(0.026)	(0.024)	(0.022)
Strong Red	0.765***	0.753***	0.747***	0.740***
	(0.021)	(0.022)	(0.020)	(0.017)
Very Strong Red	0.790***	0.790***	0.786***	0.780***
	(0.022)	(0.022)	(0.024)	(0.021)
Strong Green × SIGNALS	0.028***	0.025**	0.006	0.003
	(0.011)	(0.011)	(0.024)	(0.022)
Weak × SIGNALS	−0.111***	−0.108***	−0.120***	−0.123***
	(0.030)	(0.030)	(0.031)	(0.030)
Strong Red × SIGNALS	−0.089	−0.090	−0.116*	−0.122*
	(0.076)	(0.074)	(0.067)	(0.063)
Very Strong Red × SIGNALS	−0.066	−0.065	−0.099	−0.114*
	(0.068)	(0.068)	(0.069)	(0.067)
Game Round Fixed Effects	No	Yes	Yes	Yes
Game Fixed Effects	No	No	Yes	Yes
Participant Fixed Effects	No	No	No	Yes
Observations	44,460	44,460	44,460	44,460
Adjusted R^2	0.271	0.274	0.285	0.321

Notes: *p<0.1; **p<0.05; ***p<0.01. Clustered standard errors by session in parentheses.

G Additional Individual Results

For participant covariates we include: `female`, which is an indicator variable that takes the value of one if the participant identifies as female and zero, otherwise. `stem`, which is an indicator variable that takes the value of one if the participant's major is STEM and zero, otherwise. `overconfidence` measures the extent of a participant's over-estimation of her IQ, which is given by the difference between the number of questions a participant believes she solved correctly and the actual number of correct answers. `risk` is measured by the number of points invested in a risky asset as specified in a risky investment task solved at the end of the game.



Panel (a): Small Groups

Panel (b): Large Groups

Panel (c): Large - Small Groups

Figure G.1: Learning from others' actions in ALL treatments

Notes: Panel (a) depicts the probability of choosing red as a function of the share of red actions of other group members in the ALL4 treatment, obtained from a Bayesian logistic regression of the subject's action on the share of others' actions in the previous round conditional on the difference between red and green signals. Solid lines depict the median of the posterior distribution and dashed lines depict 95% confidence intervals. Panel (b) presents the same exercise for the ALL treatment. Panel (c) shows the difference in the probability of a red bet between the ALL and ALL4 treatments for a *Weak* signal strength.

Table G.1: Probability of Correct Guess by IQ Level (ALL versus SIGNALS)

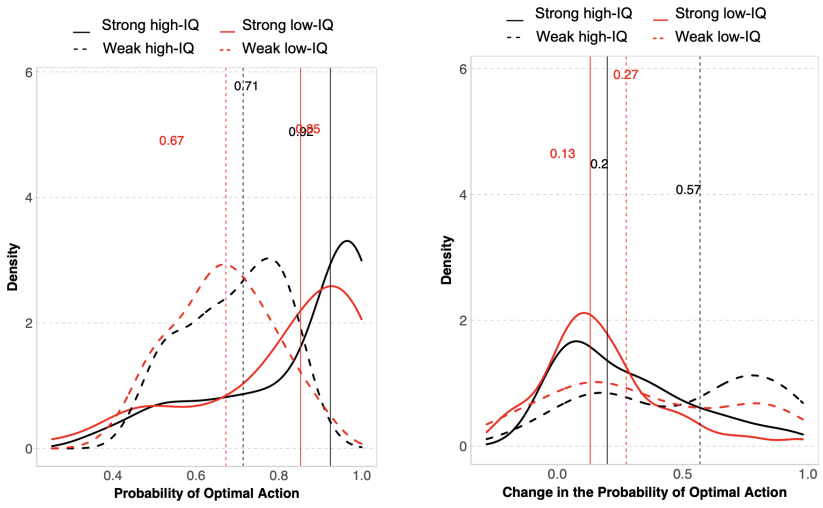
	Dependent variable:			
	Correct Guess			
	(1)	(2)	(3)	(4)
Constant	0.740***	0.614***	0.577***	0.506***
	(0.021)	(0.019)	(0.022)	(0.038)
SIGNALS	−0.055**	−0.055**	−0.055**	−0.041
	(0.026)	(0.026)	(0.026)	(0.029)
IQ	0.024***	0.024***	0.024***	0.031***
	(0.004)	(0.004)	(0.004)	(0.006)
female				0.023
				(0.022)
stem				0.018
				(0.021)
overconfidence				0.015**
				(0.007)
risk				0.0002
				(0.0001)
IQ × SIGNALS	0.003	0.003	0.003	0.001
	(0.011)	(0.011)	(0.011)	(0.010)
Game Round Fixed Effects	No	Yes	Yes	Yes
Game Fixed Effects	No	No	Yes	Yes
Participant Covariates	No	No	No	Yes
Observations	44,460	44,460	44,460	44,460
Adjusted R^2	0.014	0.032	0.034	0.037

Notes: *p<0.1; **p<0.05; ***p<0.01. Clustered standard errors by session in parentheses.

Table G.2: Probability of Correct Guess by low-IQ/high-IQ Subjects (ALL versus SIGNALS)

	Dependent variable:			
	Correct Guess			
	(1)	(2)	(3)	(4)
Constant	0.863^{***}	0.737^{***}	0.701^{***}	0.653^{***}
	(0.015)	(0.014)	(0.015)	(0.033)
SIGNALS	-0.047	-0.047	-0.047	-0.044
	(0.045)	(0.045)	(0.045)	(0.042)
low-IQ	-0.074^{***}	-0.074^{***}	-0.074^{***}	-0.088^{***}
	(0.016)	(0.016)	(0.016)	(0.026)
`female`				0.023
				(0.024)
`stem`				0.023
				(0.020)
`overconfidence`				0.009
				(0.007)
`risk`				0.0003^{*}
				(0.0001)
low-IQ × SIGNALS	0.010	0.010	0.010	0.017
	(0.031)	(0.031)	(0.031)	(0.031)
Game Round Fixed Effects	No	Yes	Yes	Yes
Game Fixed Effects	No	No	Yes	Yes
Participant Covariates	No	No	No	Yes
Observations	44,460	44,460	44,460	44,460
Adjusted R^2	0.012	0.029	0.031	0.035

Notes: *p<0.1; **p<0.05; ***p<0.01. Clustered standard errors by session in parentheses.



Panel (a): Responsiveness to signals

Panel (b): Responsiveness to actions

Figure G.2: Responsiveness to signals and actions, individual level data (ACTIONS)

Notes: Panel (a) shows the kernel distributions of participants' responsiveness to signals for both weak and strong signals. The vertical lines and the numbers next to them depict median responsiveness for each group. Responsiveness to signals is calculated based on equation (3) in the main text for ACTIONS. Responsiveness is given by the probability that a participant's action matches signal majority, i.e., the probability of an optimal action. Panel (b) shows the kernel distributions of participants' responsiveness to others' actions for weak and strong signals. The vertical lines and the numbers next to them depict median responsiveness for each group. Responsiveness to others' actions is measured by the change in the probability of choosing the action of the majority of signals when all versus none of the other group members choose the majority of signals in the last round.