Individual vs. Social Learning: An Experiment

Piotr Evdokimov* and Umberto Garfagnini†

January 30, 2020

Abstract

We design an experiment to compare individual and social learning. Player A observes private signals about a state and chooses an action that gets rewarded if and only if it identifies the state correctly. Player B observes Player A's past actions in addition to her private signals before choosing her action. The results show that while Player A's actions have a significant effect on the actions of Player B, information aggregation fails in the sense that the likelihood of choosing the correct action is the same across treatments. We argue that this is driven by the failure of Player A's actions to correctly transmit her private information. Thus, a counterfactual analysis shows that even a Bayesian Player B would not benefit from observing Player A's actions in the experiment. We also investigate the case in which both players observe each other's actions and find a similar failure of information aggregation.

JEL Classification: C92, D8

Keywords: Private signals, social learning, repeated interaction

1 Introduction

This paper studies how people learn from observing each other's actions. While this question and its implications for information aggregation have been explored both theoretically (Banerjee, 1992; Bikhchandani et al., 1992) and experimentally (Anderson and Holt, 1997), an important limitation of previous work is that agents are assumed to move sequentially. A large but separate body of literature has investigated how a long-lived decision maker processes sequential

*Department of Applied Economics, Higher School of Economics.

†School of Economics, University of Surrey.

information without observing the decisions of others (Grether, 1980, 1992). In practice, many interactions encompass features of both individual and social learning.[1] For example, Coleman et al. (1957) consider the problem of a physician adopting a new drug. The quality of the drug is uncertain and revealed through patient outcomes over time. A single doctor therefore acts as an individual decision maker updating prior beliefs on the basis of sequential information. However, the doctor can also learn about the quality of the drug by observing that other doctors adopted it. Coleman et al. (1957) show that more socially integrated doctors adopt the drug faster, a result that can be interpreted as evidence of social learning. More recently, Conley and Udry (2010) study the adoption of new agricultural technology in Ghana and find that farmers learn about its quality from their neighbors. This problem is also related to the framework considered in the present paper: the quality of the new technology is revealed to individual decision makers over time through others' choices in addition to their own.

Theoretically, the study of social learning with long-lived agents proved to be challenging. A recent contribution by Harel et al. (2018) describes a setting where information aggregation fails in the long-run, especially in a large society. This surprising result is driven by endogenous correlation in information arising from agents repeatedly observing each other's actions, which leads to a phenomenon the authors call *groupthink*—rational agents ignoring their own private information and taking identical actions for a prolonged period of time. While it provides intriguing theoretical results, the paper focuses on characterizing asymptotic behavior and has little to say about what happens in the short term; this opens a gap for an experiment to shed light on short-term behavior in the presence of both individual and social learning.

Our paper provides the first experimental study of social learning with long-lived agents. We focus on two main research questions:

QUESTION 1. *What is the effect of observing other agents' actions on behavior?*

QUESTION 2. *Does social learning improve decision making compared to individual learning?*

The experiment is designed to disentangle the effects of individual and social learning on behavior. Subjects are matched into teams of two and observe signals about an unknown binary state of the world over multiple consecutive periods. As in typical social learning experiments, subjects are incentivized to choose the action matching the unknown state.[2] The **individual**

[1]Here and below, individual learning refers to situations where decisions of other agents are not observed.

[2]To remove repeated game effects and any intertemporal hedging, each subject is paid on the basis of one randomly chosen period. Thus, subjects are incentivized to choose the myopically optimal action in every period. This protocol is standard in individual learning experiments as well as those investigating social learning on networks, which we discuss below.

learning treatment is a standard belief updating task where each subject observes only a private signal in each period. This serves as a benchmark to assess decision quality on the basis of individual learning alone.[3] Each subject in the individual learning treatment is matched with a subject in the **unilateral social learning treatment**. In every period after period one, a subject in the unilateral social learning treatment observes her partner's previously selected action in addition to her own private information. Comparing behavior in the individual learning and unilateral social learning treatments isolates the effect of observational learning on decision making. In the **bilateral social learning treatment**, both players observe each other's previous actions in every period after period one in addition to their own private information. A comparison between the unilateral and bilateral social learning treatments isolates the effect of endogenously correlated information on behavior. Intuitively, the unilateral social learning treatment simplifies the social learner's inference problem as her partner's action is based only on her private information. Correlation increases the complexity of the inference problem, which could have a detrimental effect on efficiency.[4]

We now summarize our main results:

Result 1 (The Effect of Other Players' Actions). *Observing another player's action has a significant effect on behavior in both social learning treatments.*

To establish Result 1, we identify the effect of the other player's action using that player's observed signals as instruments. We argue that because the two players' signals are conditionally independent, the instrument is valid once the observing player's information is controlled for.

Having established that observed actions of others have a significant effect on own behavior, we turn to the question of whether social learning leads to information aggregation and find that information aggregation fails:

Result 2 (Failure of Information Aggregation). *On average, belief accuracy and the propensity to make a mistake do not significantly differ across treatments.*

In case of the unilateral treatment, however, the overall lack of a treatment effect obscures a significant interaction between the effect of observing the other player's actions and the quality of the social learner's private information:

---

[3]If subjects were Bayesian, we could simply construct conterfactual choices to be used as a benchmark. However, subjects in experiments are known to process information in a conservative manner (Edwards, 1968) which can lead to mistakes, as we indeed find in the data.

[4]Enke and Zimmermann (2017) investigate subjects' ability to filter out redundant information and show a significant "correlation neglect."

Result 3 (The Effect of Social Information). *In the unilateral social learning treatment, observing your partner's action improves both decision quality and belief accuracy if your own information is bad and decreases them if your own information is good. In the bilateral social learning treatment, observing your partner's action has no significant effect on decision quality or belief accuracy.*

Result 3 suggests that social learning has less of an impact in the bilateral treatment, where the information extraction problem is more difficult.

We attribute the observed failure of information aggregation (Result 2) to shortcomings of individual learning. Because subjects in the individual treatment commit a mistake 34% of the time, compared to 16% for a Bayesian decision maker, the partner's action in both of the social learning treatments transmits too little of the partner's own information. As a result, the social learner's problem requires that the partner's mistakes are taken into account, making inferences difficult. For instance, a Bayesian decision maker in the unilateral social learning treatment who assumes her partner to be Bayesian (but uses the partner's observed non-Bayesian actions for her inferences) would commit a mistake as often as if she ignored her partner's actions altogether.

This paper contributes to the experimental literature on social learning. Following the theoretical models of Banerjee (1992) and Bikhchandani et al. (1992), Anderson and Holt (1997) find evidence of rational cascades in the laboratory. This has given rise to an extensive literature investigating the robustness of informational cascades to various conditions such as different institutional arrangements (Hung and Plott, 2001),[5] longer horizons (Goeree et al., 2007), markets (Drehmann et al., 2005; Cipriani and Guarino, 2005), costs for acquiring information (Kübler and Weizsäcker, 2004), etc.[6] The most closely related paper is Eyster et al. (2018), which investigates the effect of information redundancy arising from social learning on individual decisions. When optimal behavior requires subjects to anti-imitate their predecessors, Eyster et al. find a strong redundancy neglect that leads to inefficiencies. Our work differs from others in this literature in two important respects. First, our subjects have access to a sequence of private signals, as opposed to a single one. Second, our subjects repeatedly observe each other's actions.

Another strand of related literature has investigated social learning on networks. Choi et al. (2012) report the results of an experiment based on the theoretical model of Gale and Kariv

[5]Hung and Plott (2001) find that rewarding agents according to whether a majority of decisions is correct or incorrect leads to a significant decline in informational cascades compared to the standard setup of Anderson and Holt.

[6]Çelen and Kariv (2004) design an experiment to distinguish between informational cascades–ignoring one's own private information–and herd behavior–making identical decisions–by eliciting subjects' beliefs about the state of the world. They find that subjects do herd but most herds occur on the correct action, despite the opposite theoretical prediction.

(2003). Three subjects are matched together to form a three-player network and are incentived to choose an action to match a binary state of the world. Subjects observe a signal about the state only in the first period and choose actions over multiple periods, with each subject observing the choices of others in the network. In a similar framework Chandrasekhar et al. (2019) estimate the proportion of Bayesian and DeGroot learning types in two laboratory populations and find a non-negligible share of DeGroot types.[7] While subjects in these papers also make repeated choices, they only observe a single signal at the beginning of the game as opposed to a sequence of private signals.

2 Experimental Design

At the beginning of the experiment subjects are matched in teams of two. Within a team, each subject is randomly assigned to one of two roles: Player A or Player B, with exactly one subject in each role. The roles and teams stay fixed for the duration of the experiment. A session lasts for 20 incentivized periods.

Subjects are told that there are two urns, *orange* and *purple*, each containing three balls. The orange urn contains two orange balls and one purple ball, while the purple urn contains one orange ball and two purple balls. At the beginning of the experiment, the computer selects one of the two urns with equal probability for each two-player team. None of the subjects are told which urn is selected for their team. In each period, each subject privately observes a ball drawn with replacement from the selected urn, which is her private signal. In each period, every subject also faces two tasks: the *matching task* and the *guessing task*. In the matching task, the subject must choose an action, orange or purple. In the guessing task, the subject is asked to provide an estimate of the likelihood that the secret urn is orange. Both tasks are incentivized, as explained below.[8]

Our experimental design consists of three treatments. In the **individual treatment**, the subject in the role of Player A only observes her selected ball in each period of the experiment, as in standard belief updating tasks. For each team, the subject in the role of Player B is assigned to a condition we call the **unilateral social learning treatment**. In every period after period one, this player also observes her partner's previously selected action.[9] In the **bilateral**

[7]A DeGroot type chooses an action which conforms with the majority in the previous period. See also Mueller-Frank and Neri (2013) for another experiment on social learning in networks.

[8]We elicit a subject's beliefs about the state of the world, which we use in our estimation analysis, to shed light on the underlying decision making process (as in Hung and Dominitz, 2004; Drehmann et al., 2007).

[9]The experimental interface contains a table which lists the history of balls observed by a subject up to and including the current period, as well as the history of actions chosen by her partner up to and including the

social learning treatment, both subjects in the same team observe each other's previous actions in every period after period one.[10] While there are no direct payoff externalities in any of our treatments, subjects in the two social learning conditions are affected by informational externalities through observability of their partner's past actions.

Thus, the individual learning treatment is a standard updating task. The unilateral social learning treatment is a standard updating task augmented by observations of another player's actions and knowledge that the other player does *not* observe your own actions.[11] The bilateral social learning treatment is a standard updating task augmented by observations of another player's actions and knowledge that the other player observes your own actions.

Incentives. Each subject is paid on the basis of one randomly chosen decision among her matching tasks, and one randomly (and independently) chosen decision among her guessing tasks. Paying for only one randomly chosen decision for each task is meant to minimize any intertemporal hedging across periods and decisions, making each choice as close as possible to a standalone decision problem.[12] A subject earns \$2.50 if her action matches the color of the selected urn and \$0.80 otherwise. The subject's response to the guessing task determines the likelihood of receiving a bonus of \$0.50 or \$0.20 according to the Binarized Scoring Rule of Hossain and Okui (2013).[13] The elicitation procedure is incentive compatible irrespective of attitudes toward risk and relatively simple to implement. The rule works as follows: *i*) the decision maker makes a report $r \in [0,1]$;[14] *ii*) given the state of the world θ, the decision maker's loss is computed according to the quadratic scoring rule, that is, $L(r,\theta) = \left(r - \mathbf{1}_{\{\theta=Orange\}}\right)^2$;

previous period (as in Chandrasekhar et al., 2018). A screenshot is available in the online appendix, which can be found at https://drive.google.com/open?id=1F-c4tE6snYpErnYboq1BOjwWDbRaEvLF.

[10] Recall that each subject also observes a private signal about the state of the world regardless of the treatment.

[11] The instructions explain that there are two player roles, Player A and Player B. Player A only observes her own sequence of private signals, while Player B observes both her own sequence of private signals and Player A's past actions. Understanding of the difference between the two player roles is tested in the quiz. Once players are matched and informed of their roles, the screen reminds them again about the difference between the two roles. Subjects are also reminded of their role on each decision screen. Subjects in the role of Player A are reminded that they cannot observe the past actions of Player B, while subjects in the role of Player B are reminded that they can observe Player A's actions whereas Player A cannot observe their own actions.

[12] Azrieli et al. (2018) show theoretically that selecting one task at random is the only incentive compatible way to pay subjects under a mild monotonicity assumption on subjects' preferences. Our procedure is common to other recent experiments (e.g. Charness et al., 2018).

[13] The bonuses from the guessing task are significantly smaller than those for the matching task to have subjects focus their attention primarily on the matching task (as in, e.g., Charness et al., 2018). Moreover, the bonuses for incorrect decisions are designed to guarantee a minimum payment of \$1.00 to participating subjects. Thus, the bonuses act as a guaranteed minimum payment which subjects receive only if they complete the experiment. This is extremely important for online experiments, where subject dropouts are a serious concern which can be reduced through guaranteed payments.

[14] The belief set in the experiment was a discretized version of [0,1] with $R = \{0, 0.01, 0.02, ..., 0.98, 0.99, 1\}$.

iii) finally, the computer draws a number k uniformly at random from the interval $[0, 1]$. If $L(r, \theta) \leq k$, the decision maker earns \$0.50; otherwise, she earns \$0.20.[15]

As is standard practice in belief updating and social learning experiments, subjects receive no feedback about their own performance or the performance of their matched partner during the experiment. At the end of the experiment, subjects receive information about which urn was selected, which two decisions have been randomly chosen for payment, which number has been drawn to implement the Binarized Scoring Rule for the guessing task, and their bonuses.[16]

2.1 Predictions

In the unilateral social learning treatment, observing the other player's action provides an additional source of information about the state of the world and should improve the quality of decisions and the accuracy of beliefs. To gain some intuition, suppose that Player A's signals are publicly observed in every period, then the probability that Player B (in the unilateral social learning treatment) makes a mistake in the matching task is given by:[17]

$$\text{Prob(Mistake in period } t) = e^{-2\eta t + o(t)} \tag{1}$$

where $\eta \equiv -\inf_{x\geq 0} \ln E_{Orange}\left[e^{-xl}\right] = -\inf_{x\geq 0} \ln E_{Purple}\left[e^{-xl}\right]$, and l is the log-likelihood ratio. This means that the probability of making a mistake in period t when observing two conditionally independent signals in every period decays exponentially at a rate equal to 2η. By contrast, the same probability in the individual learning treatment decreases at rate η. Thus, if Player B in the unilateral social learning treatment observed her partner's signals, she would learn asymptotically faster than her partner. In the experiment, Player B can only observe her partner's past action, not her signal. This ought to slow down her rate of learning but, provided that Player A's actions are informative about her private signals, the rate of learning would still be faster than that of Player A. This motivates the following prediction:

Prediction 1. *Subjects are more likely to choose correct actions in the unilateral social learning treatment than in the individual learning treatment.*

[15]With this elicitation procedure, a subject's optimal response is to maximize the probability of getting the higher prize which corresponds to reporting the probability that the selected urn is orange, regardless of attitudes toward risk.The payment rule is clearly explained to the subjects in the instructions and several examples are provided, including a table indicating the loss corresponding to one's decision conditional on each state of the world. Each decision screen further summarizes the most important information from the instructions relevant for the current decision, including a table of potential losses. The instructions and screenshots can be found in the online appendix.

[16]More detailed information about the implementation of the experiment can be found in Appendix A.

[17]This is a well-known result (e.g., see Harel et al, 2018).

In the bilateral social learning treatment, the player's partner observes the player's previous actions. Thus, the partner's action contains redundant information that is correlated with the player's own information. This endogenous correlation between agents' actions can lead to inefficiencies, as shown theoretically by Harel et al. (2018).[18] Moreover, recent experiments have shown that people have difficulty dealing with correlated information (see, Enke and Zimmermann, 2017; Eyster and Weizsäcker, 2016; Eyster et al., 2018), which might reduce decision quality further. This motivates the following prediction for the bilateral social learning treatment:

PREDICTION 2. *Subjects are more likely to choose correct actions in the unilateral social learning treatment than in the bilateral social learning treatment.*

3 Results

Data was collected in October 2019 from Amazon MTurk using the software oTree (Chen et al., 2016).[19] Appendix A describes the implementation details. A total of 496 subjects were successfully matched and completed the experiment,[20] with 162 subjects in the bilateral social learning treatment and 167 in each of the unilateral social learning and individual learning treatments.[21] The average hourly wage for subjects that completed the experiment is \$11.86,[22] and the aver-

[18]The results of Harel et al. (2018) focus on asymptotic behavior and the effect of team size and therefore are not directly applicable to the experiment.

[19]We reposted the MTurk HIT (Human Intelligence Task) for the bilateral social learning treatment 11 times and the combined HIT for the unilateral social learning and individual learning treatments 24 times. Each time a HIT was posted, it allowed 20 assignments to be submitted. The choice of only 20 assignments per session is because subjects tend to sign up as soon as the HIT is posted and the arrival rate of new subjects decreases with time since posting which increases the chance of unmatched participants. Amazon MTurk allows requesters (i.e., experimenters) to assign a qualification to each subject that accepts one of the HITs. This qualification prevents an MTurk worker to participate in the experiment more than once.

[20]Matching rates are notoriously low and dropout rates high in online experiments based on MTurk participants (Arechar et al., 2018). Our experiment is not immune to this issue. Several subjects accepted to participate and then withdrew from the experiment at the instructions stage. However, the vast majority of subjects that completed the quiz and decided to proceed with the experiment, if successfully matched, also completed the experiment. Some subjects could not be matched, or had to quit the experiment because their matched partner had dropped out.

[21]We also conducted a preliminary pilot session of the bilateral social learning treatment on October 1st, 2019 which collected data from 14 participants. Since we modified the instructions and the structure of the experiment after this session, we do not include the data from the pilot in the analysis.

[22]Hara et al. (2018) estimate an average hourly wage per worker on Amazon MTurk of between \$3.13 and \$3.48. Thus, the average compensation in our experiment is close to 4 times the average wage as well as above the minimum federal wage of \$7.25, and very close to the California minimum wage of \$12/hour. Furthermore, the median hourly wage per worker is estimated to be between \$1.77 and \$2.11, compared to a median hourly wage per worker of \$11.52 from participating in our experiment.

age subject spent 17 minutes to complete the experiment.[23] Figure 1 shows the distribution of pre-experiment quiz scores for all the subjects that participated and completed the experiment. The median quiz score is 5 out 6 in each treatment (see Table 5 in the Appendix), which is encouragingly high for an online experiment.

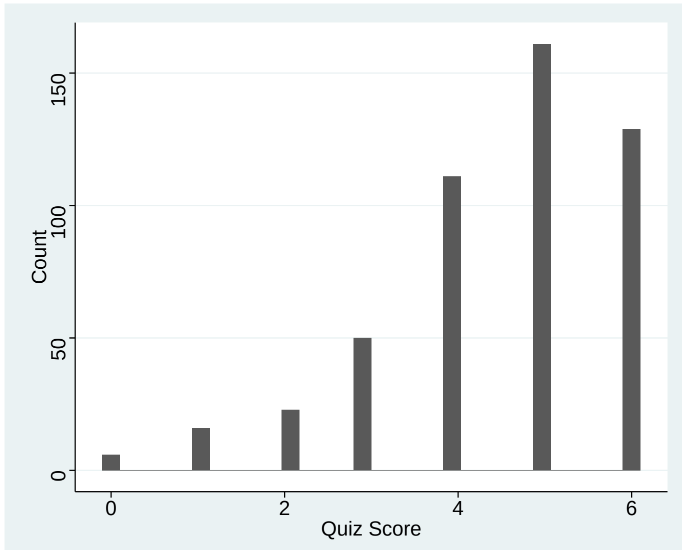


Figure 1: Distribution of quiz scores.

Our analysis proceeds as follows. We first show that observing other players' previous actions has an effect on behavior using a restricted subset of observations. Second, we use an IV analysis to show that this result generalize to all observations in the experiment. Having established that social learning had a significant effect on behavior, we turn to the question of whether or not information aggregation was successful and compare decision quality across the three treatments.

3.1 The effect of other players' actions

To estimate the effect of observing other players' actions, we first focus on observations where subjects privately observe equal numbers of orange and purple balls. In this case, there is no evidence favoring either the orange or the purple action if information regarding the other player's action is not received or ignored. If this information is paid attention to, we hypothesize an orange action to be more likely if the partner's previous action was orange. To test this hypothesis, we regress a subject's choice on the partner's previous choice for the aforementioned

[23]As subjects in the experiment arrive randomly over time, we report the average time spent by a subject on the experiment from the time that the assignment is accepted until it is submitted. Figure 6 in the Appendix shows the distribution of completion times for all subjects.

subset of observations. The results are reported in the first three columns of Table 1. Consistent with the hypothesis, the partner's lagged action has a significant effect in the social learning treatments ($P < 0.001$ in the unilateral and $P < 0.05$ in the bilateral treatment), but not in the individual learning treatment where the other player's action is not observed ($P = 0.669$).

While this analysis is instructive, only 876/9424 (approx. 9%) of the observations are associated with equal numbers of orange and purple balls. To generalize these findings to all periods greater than one, we regress a player's action ($OrangeChoice_{i,t}$) on a vector of controls $\mathbf{X}_{i,t}$ and the lagged action of the partner ($OrangeChoice_{-i,t-1}$). Formally, the model we estimate is:

$$OrangeChoice_{i,t} = \beta_0 + \boldsymbol{\beta}_1 \mathbf{X}_{i,t} + \beta_2 OrangeChoice_{-i,t-1} + \epsilon_{it}, \quad (2)$$

where $\mathbf{X}_{i,t}$ is a vector of controls described below. As in the case of the model in Table 1, we estimate this model for both the individual and the social learning treatments. The results for the individual treatment, where no effect of the partner's action can plausibly be observed, should be viewed as a check of the validity of our estimation exercise.

We now describe the variables contained in $\mathbf{X}_{i,t}$. First, $\mathbf{X}_{i,t}$ controls for the information observed by player i using the variable $OrangeFrac_{i,t}$, the observed fraction of orange balls in period t. This variable provides crucial decision-relevant information. In the individual treatment, for instance, a rational player would choose the orange action if the fraction of orange balls exceeds 1/2 and purple if it does not.[24]

	(1) Individual	(2) Unilateral	(3) Bilateral
Other player's action in $t-1$	-0.0244	0.229****	0.170**
	(0.0568)	(0.0594)	(0.0686)
Constant	0.514****	0.365****	0.458****
	(0.0442)	(0.0376)	(0.0494)
Observations	287	326	263

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 1: Effect of lagged partner's action on your own action in periods where equal numbers of orange and purple balls are observed.

The controls also include the variable $Orange_{i,t}$, an indicator for whether the most recent observed ball was orange. As we discuss below, the observation that $Orange_{i,t}$ has a significant effect even when $OrangeFrac_{i,t}$ is controlled for may capture a "recency effect" which has been documented in the literature (e.g., see Tubbs et al., 1993). We also show in the appendix that

[24]Recall that the subject sees her private signal history on the screen before making each decision.

	(1)	(2)	(3)	(4)	(5)	(6)
		OLS			IV	
	Individual	Unilateral	Bilateral	Individual	Unilateral	Bilateral
Other player's action in $t-1$	0.0241	**0.104******	**0.0683******	0.00303	**0.179*****	**0.135****
	(0.0158)	(0.0207)	(0.0189)	(0.0661)	(0.0642)	(0.0619)
Fraction of orange balls in t	0.224***	0.0686	0.236****	0.240***	-0.0249	0.0934
	(0.0702)	(0.0661)	(0.0612)	(0.0915)	(0.0877)	(0.0789)
Orange ball in t	0.237****	0.278****	0.240****	0.223****	0.273****	0.254****
	(0.0398)	(0.0322)	(0.0357)	(0.0400)	(0.0338)	(0.0361)
Fraction of orange actions in $t-1$	0.588****	0.539****	0.495****	0.730****	0.720****	0.667****
	(0.0385)	(0.0399)	(0.0398)	(0.0447)	(0.0508)	(0.0524)
Constant	-0.0448**	0.0112	0.00610	-0.104****	-0.0658**	-0.0483*
	(0.0223)	(0.0284)	(0.0248)	(0.0266)	(0.0311)	(0.0293)
Observations	3173	3173	3078	2505	2505	2430

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 2: The effect of other players' actions.

introducing the lagged indicators $Orange_{i,t-1}$, $Orange_{i,t-2}$, etc., as additional controls leaves our main conclusions unaffected.

Crucially, $\mathbf{X}_{i,t}$ controls for player i's previous actions. The motivation for this is as follows. $OrangeChoice_{-i,t-1}$ is made by player $-i$ after $-i$ observes the actions $OrangeChoice_{i,t-2}$, $OrangeChoice_{i,t-3}$, etc., of player i. Because the period-t action of player i is likely influenced by player i's actions in the previous periods (this will be true if there is any serial correlation in behavior, for example), this introduces an endogeneity issue in the model described by Equation (2). To control for player i's lagged actions, we include the variable $FracOrangeChoice_{i,t-1}$ in $\mathbf{X}_{i,t}$. This variable represents the fraction of orange choices made by player i in periods up to period $t-1$. Robustness checks in the appendix also control for player i's lagged actions $OrangeChoice_{i,t-1}$, $OrangeChoice_{i,t-2}$, etc., directly. The robustness checks have no substantive impact on our main conclusions, providing further evidence that these conclusions are robust to functional form assumptions.

The OLS results are reported in the first three columns of Table 2. The effect of $OrangeChoice_{-i,t-1}$, the action of the other player, is significant in the two social learning treatments ($P < 0.001$) but not significant in the individual learning treatment ($P = 0.128$). We highlight this result as follows:

Result 1 (The Effect of Other Players' Actions). *Observing another player's action has a significant effect on behavior in both social learning treatments.*

As mentioned above, we provide several robustness checks in the appendix. The first three columns of Table 6 use probit instead of OLS regressions; the first three columns of Table 7 control for lagged orange balls observed by player i in addition to the overall fraction; the first three columns of Table 8 control for player i's lagged actions. All of the specifications show the effect of the partner's lagged action to be significant in the two social learning treatments but insignificant in the individual treatment, providing additional evidence of Result 1.

It is natural to ask if the partner's action in period $t-2$ has an effect on period t behavior in the two social learning treatments. The first three columns of Table 9 in the appendix add the variable $OrangeChoice_{-i,t-2}$ to the specification in Equation (2) and find its effect to be insignificant in the two social learning treatments ($P = 0.497$ in the unilateral and $P = 0.147$ in the bilateral treatment). Thus, only the most recently observed action of one's partner has an impact on behavior.

3.2 The instrumental variables approach

While the analysis in the previous section deals with endogeneity of the other player's action by introducing a number of controls, our experimental design also allows us to use an instrumental variables approach. As instruments for $OrangeChoice_{-i,t-1}$, we use the variables $Orange_{-i,t-k}$ for $1 \leq k \leq 5$, the balls observed by player $-i$ in periods $t-1$, $t-2$, and so on. Note that this requires us to restrict our attention to observations in period 6 and above.[25]

To see that the instruments are valid, consider that $-i$'s previously observed balls are determined solely by the color of the urn and chance. If the urn is orange (resp., purple), each ball is more likely to be orange (resp., purple). The urn also determines the information player i observed in the past. Our claim is that the variables included in $\mathbf{X}_{i,t}$ completely capture the effect of the decision-relevant information generated by the urn for player i,[26] making the instruments valid.

The results for the baseline linear model are reported in the last three columns of Table 2. Following the approach of the previous section, the appendix includes a number of robustness checks. The last three columns of Table 6 use IV probit regressions; the last three columns of Table 7 control for lagged orange balls observed by player i; the last three columns of Table 8 control for player i's lagged actions. Consistent with Result 1, all of the specifications show the effect of the partner's lagged action to be significant in the two social learning treatments but

[25] Expanding and contrasting the set of instruments has no substantive impact on our main conclusions.

[26] In particular, note that the effect of the indicator variables $Orange_{i,t-1}$, $Orange_{i,t-2}$, etc. is not significant if these variables are introduced in the regression (Table 7). This provides supporting evidence for the claim that the controls in $\mathbf{X}_{i,t}$ capture all decision-relevant information generated by the urn.

insignificant in the individual treatment.

The first-stage F statistics are high in all of our specifications. Thus, in the baseline linear model, the Kleibergen-Paap F statistic is equal to 20.486 in the unilateral and 26.164 in the bilateral social learning treatment. Both of these statistics exceed the 5% bias threshold relative to OLS according to the Stock-Yogo test of weak identification. Because the model is over-identified, we can also test for the validity of the instruments. Using the Hansen J statistic, in the baseline linear model, the null hypothesis of valid instruments cannot be rejected in either the unilateral ($P = 0.2076$) or bilateral ($P = 0.7607$) social learning treatment. Similar results are observed in the robustness checks.

Taken together, the results in Sections 3.1 and 3.2 suggest that when observed, the lagged action of one's partner had a significant effect on behavior. We now turn to the question of whether this effect led to less mistakes and more accurate guesses relative to the individual learning treatment, i.e. if information aggregation was successful.

3.3 Information Aggregation

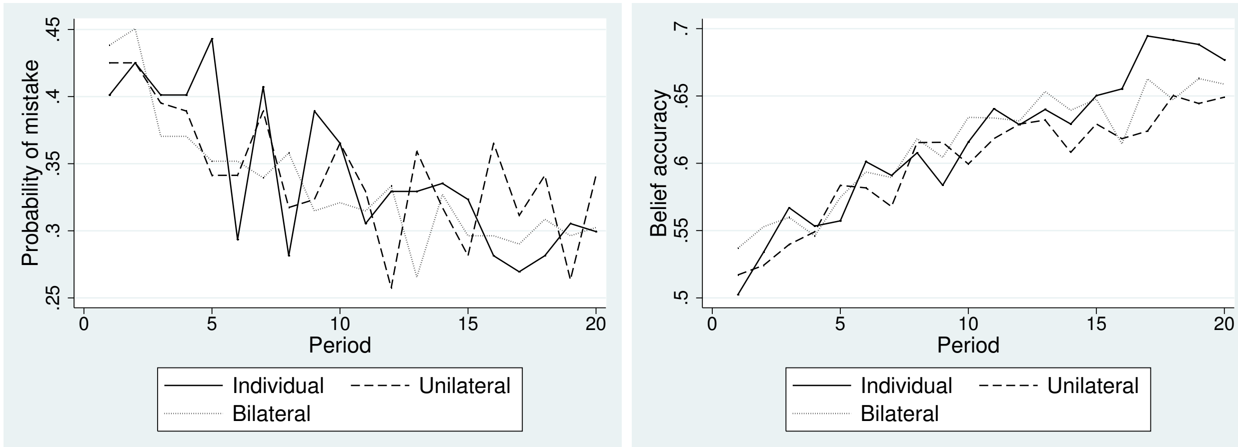
Observing the other player's actions has an effect on behavior, but does it help? Figure 2 reports the probability of making a mistake (a) and belief accuracy (b) in each period of each treatment.[27] The figure suggests that despite the fact that subjects were affected by the actions of their partners, observing other players' actions did not improve decisions or beliefs on average in either of the social learning treatments. We test the significance of these findings in Table 3.[28] The first column, which regresses a dummy variable for making a mistake against treatment dummies for the unilateral and bilateral social learning treatments, shows that neither dummy variable is significant (smallest $P = 0.693$). The same is true in the second column, where the dependent variable is belief accuracy (smallest $P = 0.402$). We highlight these results as follows:

Result 2 (Failure of Information Aggregation). *On average, belief accuracy and the propensity to make a mistake do not significantly differ across treatments.*

We now show that the lack of average treatment effects obscures significant interactions between treatment effects and the quality of private information. To this end, we define a variable $TrueFrac_{it}$ as the fraction of the correctly-colored balls observed by subject i in period t. Say that information is of low quality if $TrueFrac_{it} < 0.5$ and high quality if $TrueFrac_{it} > 0.5$. In

[27] Belief accuracy is measured as the reported belief normalized by the correct value of the state, i.e. as x when the urn is orange and $1 - x$ when the urn is purple, where x is the reported belief.

[28] We exclude period one observations from the analysis.



(a) Probability of making a mistake. (b) Belief accuracy.

Figure 2: Mistakes and belief accuracy by treatment.

	(1) Mistake	(2) Accuracy	(3) Mistake	(4) Accuracy
Unilateral social learning	-0.000630	-0.0172	-0.224***	0.132***
	(0.0270)	(0.0205)	(0.0711)	(0.0506)
Bilateral social learning	-0.0109	-0.00427	-0.0708	0.0645
	(0.0277)	(0.0219)	(0.0680)	(0.0576)
$TrueFrac_{it}$			-0.863****	0.565****
			(0.0692)	(0.0592)
Unilateral social learning × $TrueFrac_{it}$			0.352***	-0.235***
			(0.110)	(0.0818)
Bilateral social learning × $TrueFrac_{it}$			0.126	-0.126
			(0.0994)	(0.0932)
Constant	0.340****	0.621****	0.899****	0.256****
	(0.0204)	(0.0155)	(0.0445)	(0.0350)
Observations	9424	9424	9424	9424

Subject-clustered standard errors in parentheses
$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$, $**** \ p < 0.001$

Table 3: Treatment effects.

our data, the likelihood of holding low quality information in an arbitrary period varies between 19% in the bilateral and 24% in the unilateral social learning treatment.

When $TrueFrac_{it} \leq 0.5$, the subject's own private information points to the incorrect action and mistakes are more likely. In this case, because the subject's partner is more likely to have high than low quality information, placing a positive weight on the partner's past action might improve decision quality.[29] On the other hand, if $TrueFrac_{it}$ is abnormally high, e.g. close to one, then placing a positive weight on the partner's past action might decrease decision quality. Figure 7 in the appendix reports the probability of making a mistake in each period conditional on the quality of one's information.

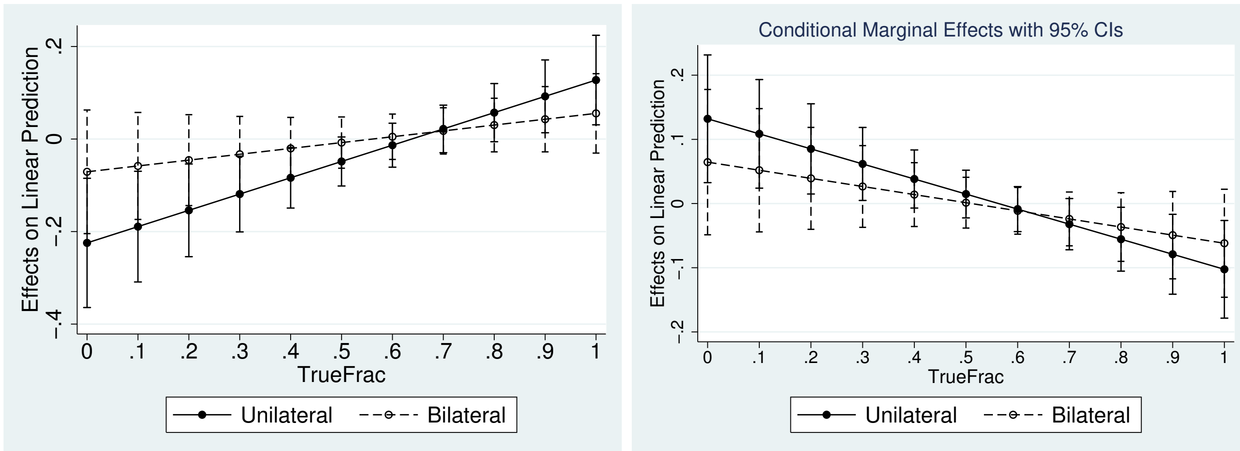
We re-run the regressions in the first two columns of Table 3 controlling for this variable as well as its interactions with both treatment dummies and report the results in the last two columns of Table 3. For ease of interpretation, Figure 3 reports the marginal effects of both treatment dummies on the probability of making a mistake (a) and belief accuracy (b) for each value of the $TrueFrac_{it}$ variable, together with 95% confidence intervals.

The marginal effects in the unilateral treatment are shown as solid lines. These results are aligned with the intuition of the previous paragraph. When $TrueFrac_{it}$ is low, the probability of making a mistake is significantly lower in the unilateral than individual treatment ($P < 0.01$ when $TrueFrac_{it} = 0$), and when $TrueFrac_{it}$ is close to one, the probability of making a mistake is higher ($P < 0.01$ when $TrueFrac_{it} = 1$). Similarly, the subjects' guesses are more accurate in the unilateral than individual treatment when $TrueFrac_{it}$ is low ($P < 0.01$ when $TrueFrac_{it} = 0$) and less accurate when $TrueFrac_{it}$ is high ($P < 0.01$ when $TrueFrac_{it} = 1$).

The marginal effects in the bilateral treatment are shown as dashed lines. Although the effects are of the same sign as those in the unilateral treatment, they never reach significance (smallest $P = 0.149$). This result is intuitive: because one's own actions have an influence on the action of one's partner (as shown in Section 3.1), one's incorrect information will be incorporated into (incorrect) actions of your partner. Thus, the smaller marginal effects in the bilateral compared to the unilateral treatment point to the information extraction problem being more difficult in the former case.

Result 3 (The Effect of Social Information). *In the unilateral social learning treatment, observing your partner's action improves both decision quality and belief accuracy if your own information is bad and decreases them if your own information is good. In the bilateral social learning treatment, observing your partner's action has no significant effect on decision quality or belief accuracy.*

[29]Note that the subject need not be aware of the quality of her own information for this argument to hold.



(a) Probability of making a mistake.

(b) Belief accuracy.

Figure 3: Treatment effects conditional on information quality.

Note that Result 3 also provides insight into Result 2: because the marginal effect of being in a social learning treatment is beneficial when one's information is bad and detrimental when it is good, the two forces cancel each other out when the overall marginal effect is computed.

3.4 Understanding the failure in information aggregation

Result 2 points to a failure of information aggregation in the unilateral social learning treatment, where a subject (in the role of Player B) observes draws from the urn as well as past actions of her partner, while the partner (in the role of Player A) observes only draws from the urn. Theoretically, information aggregation requires that the following two conditions are satisfied:

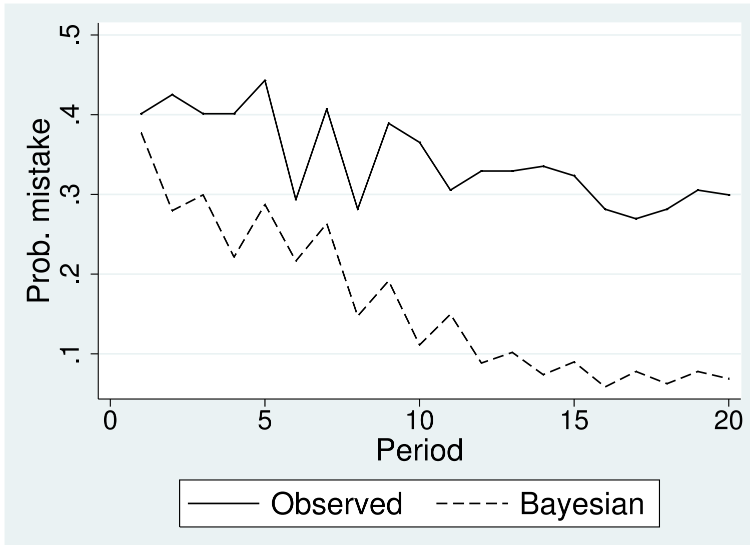
(i) Player A's action correctly transmits Player A's private information, and

(ii) Player B draws a correct inference upon observing Player A's action.

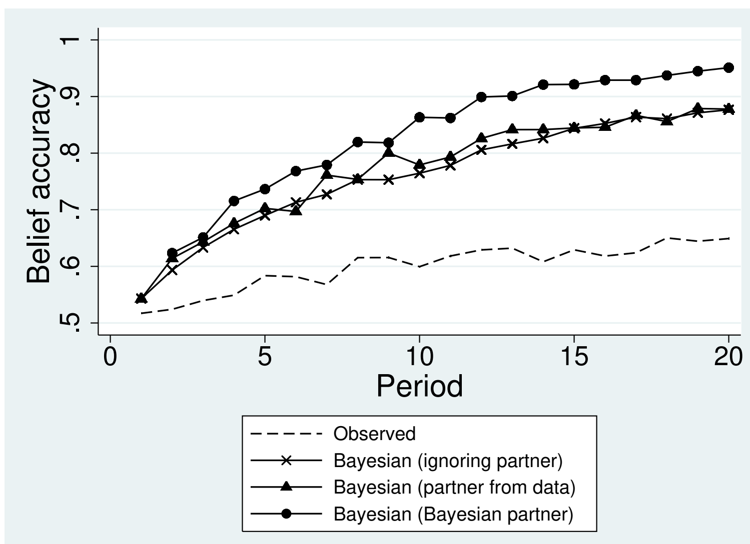
We now argue that condition (i) fails in our data. Moreover, the extent to which it fails is sufficient to explain the observed failure of information aggregation. Thus, even if Player B were Bayesian (and assumed she was facing a Bayesian partner), information aggregation in the experiment would be impossible given Player A's observed action choices.

To show that condition (i) fails, we first focus on Player A's actions in the individual treatment.[30] Figure 4 (a) compares the Bayesian benchmark for the probability of making a mistake

[30]To compute the benchmark probabilities, we first compute the action of a Bayesian decision maker for each observation. A Bayesian decision maker chooses orange (resp., purple) if she assigns a probability > 0.5 (resp., < 0.5) to the orange state. Observations in which the Bayesian posterior is equal to 0.5 are excluded from this



(a) Probability of making a mistake in the individual treatment.



(b) Belief accuracy in the unilateral treatment.

Figure 4: Average number of mistakes by period and treatment.

to that actually observed in the data. Overall, a Bayesian decision maker makes a mistake 16% of the time, while subjects in the experiment made mistakes about twice as often (34%). This shows that the amount of information transmitted by Player A's actions fell substantially short of the Bayesian benchmark. We now turn to the question of whether the degree to which it failed can by itself explain the lack of information aggregation in the unilateral treatment.

Figure 4 (b) shows the observed belief accuracy in the experiment (dashed line) together with several "Bayesian" benchmarks:

- The line with triangles indicates the beliefs of a Bayesian Player B who assumes she faces a Bayesian Player A and infers Player A's information from her actions.[31] Crucially, the *actual (non-Bayesian) actions of Player A* are used to compute Player B's counterfactual beliefs in this case.

- The line with circles is similar to the line with triangles with the crucial difference that Player B's partner is forced to be Bayesian in this case. Thus, the *actions of a Bayesian Player A* are used to compute Player B's counterfactual beliefs.

- The line with crosses indicates the beliefs of a Bayesian Player B that ignores Player A's actions (i.e., a Player B that acts as Player A).

The figure suggests that a Bayesian decision maker that ignores her partner's actions would do as well in the experiment as a Bayesian decision maker that assumes she is facing a Bayesian partner and infers her partner's information from Player A's actual actions in the experiment. I.e., no treatment effect of a unilateral vs. individual treatment would be observed on beliefs (or, by extension, the probability of making a mistake) even if Player B were Bayesian, given Players A's actions in the data. Thus, the degree to which Player A's actions fail to transmit Player A's private information is sufficient to explain Result 2.

To shed additional light on this, we run the regressions used to generate Figure 3 with counterfactual instead of observed beliefs. In the first column of Table 4, the counterfactual belief accuracy of a Bayesian decision maker is regressed against an indicator variable for being in

analysis. The Bayesian benchmark for the probability of making a mistake is simply the probability of the action of a Bayesian decision not matching the state.

[31] The details of the construction can be found in Appendix D. In a sequential setting, it is in principle possible for a Bayesian agent to infer more about her partner's private history from observing the sequence of past actions as opposed to only a single action. For example, if an agent in the individual treatment takes the orange action in an arbitrary past period n' and then switches to the purple action in period $n'+1$, then the only possibility is that she has observed an equal number of balls of each color in period n', etc. As these inferences become more complex over time, we abstract from them in the analysis and assume that an agent in the unilateral treatment only bases her current decision on the information inferred from the partner's previous action.

the unilateral treatment.[32] In the individual treatment, the counterfactual beliefs are computed using Bayes rule on the basis of Player A's observed balls. In the unilateral treatment, the counterfactual beliefs are computed using Player B's observed balls as wells as the counterfactual actions of a *Bayesian* Player A. I.e., for observations in the unilateral treatment, the dependent variable in the first column of the table corresponds to the line with circles in Figure 4 (b). We find that compared to the counterfactual decision maker in the individual learning treatment, beliefs of the counterfactual Bayesian decision maker in the unilateral treatment are 7% more accurate ($P < 0.001$). I.e, the counterfactual Bayesian beliefs in this case show evidence of information aggregation.

	(1) Bayesian partner	(2) Actual partner	(3) Bayesian partner	(4) Actual partner
Unilateral social learning	0.0700****	0.0136	0.289****	0.0848*
	(0.0196)	(0.0218)	(0.0526)	(0.0510)
$TrueFrac_{it}$			1.176****	1.176****
			(0.0444)	(0.0444)
Unilateral social learning × $TrueFrac_{it}$			-0.348****	-0.123*
			(0.0721)	(0.0692)
Constant	0.771****	0.771****	0.00917	0.00917
	(0.0155)	(0.0155)	(0.0313)	(0.0313)
Observations	6346	6346	6346	6346

Subject-clustered standard errors in parentheses
$* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$, $**** \ p < 0.001$

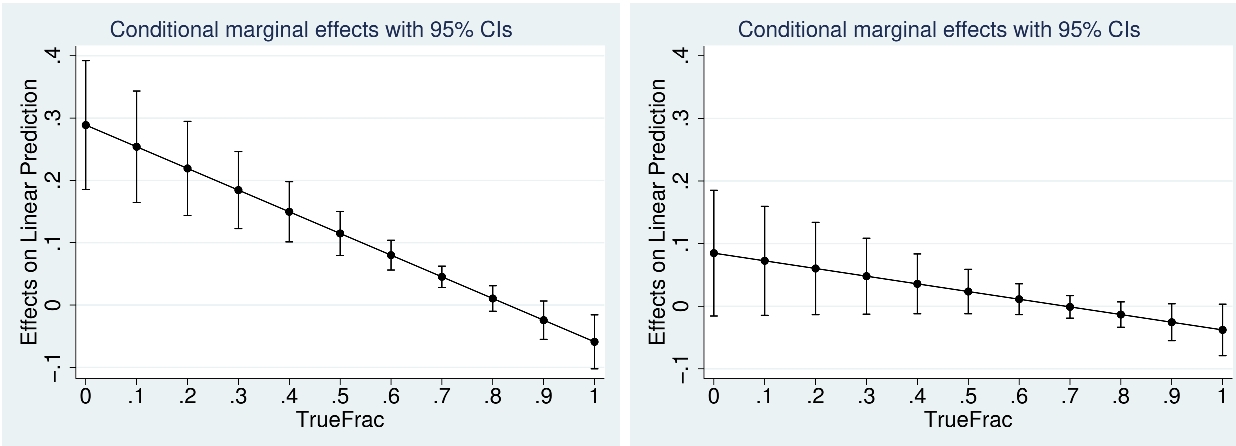
Table 4: Treatment effects on belief accuracy of counterfactual beliefs.

The regression in the second column of Table 4 is identical to that in the first column with the exception that the counterfactual Bayesian beliefs in the unilateral treatment are based on Player A's actual actions in the experiment. I.e., the dependent variable in the second column corresponds to the line with triangles in Figure 4 (b). Here, compared to the individual learning treatment, beliefs in the unilateral treatment are only 1% more accurate and the effect is not significant ($P = 0.534$). Bayesian inferences do not lead to information aggregation on average given Player A's actions in the experiment.

The third and fourth columns of Table 4 repeat the analysis in the first and second columns controlling for the quality of the counterfactual decision maker's information ($TrueFrac_{it}$) and the interaction of this variable with the treatment dummy. I.e., the analysis is similar to that

[32]The indicator variable takes on the value zero for observations in the individual treatment, and one in the the unilateral social learning treatment. The bilateral treatment is excluded from this analysis.

reported in the last column of Table 3, with the exception that the bilateral treatment is not considered and counterfactual instead of actual beliefs are used.



(a) Bayesian partner.

(b) Actual partner.

Figure 5: Treatment effects on counterfactual beliefs conditional on information quality.

For ease of interpretation, the marginal effects of observing social information for the two counterfactual decision makers are plotted in Figure 5. When Player A is Bayesian, the marginal effect on counterfactual beliefs of observing the partner's actions is large and positive for most values of the $TrueFrac_{it}$ variable (a).[33] When Player A's choices are taken from the data, the marginal effects on Player B's counterfactual beliefs become substantially smaller (b). Moreover, the marginal effects on Bayesian counterfactuals in Figure 5 (b) are remarkably close to their counterparts in the data (Figure 3 (b)). Thus, subjects in the role of Player B on average derive as much information from Player A's actions as a Bayesian decision maker would.

4 Conclusion

This paper contributes to the social learning literature by investigating information aggregation with long-lived agents. Our results show that information aggregation fails even when social learning is unaffected by endogenous correlation arising from bilateral observation of each other's actions. We attribute the failure in information aggregation to a failure of individual learning: mistakes in processing private information propagate and prevent information from being aggregated over time.

[33]Note that when one's information is remarkably good ($TrueFrac_{it} \approx 1$), the marginal effect on counterfactual Bayesian beliefs is negative, as is the case in Player B's actual reported beliefs (Figure 3).

If the observed actions are driven by mistaken beliefs, as they are in the experiment, the information extraction problem is more difficult for the social learner. We argue in the preceding section that the failure of the observed player to transmit her information is sufficient for explaining the observed failure of information aggregation. On the other hand, we cannot rule out the hypothesis that the observing player fails to extract all of the information contained in the observed player's action. Specifically, if Player A makes mistakes and Player B has a correct model of how the mistakes are made, Player B should in principle be able to extract more information from Player A's actions than a Bayesian Player B, who assumes no mistakes.

We conclude with two promising venues for future work. First, though our study restricts attention to teams of two subjects, endogenous correlation in actions grows larger as the group size increases. This is predicted to exacerbate the *groupthink* effect highlighted by Harel et al. (2018). Second, with more than two subjects in a team, the network structure can play an important role for information aggregation (Choi et al., 2012). Both of these issues deserve further investigation in the laboratory.

References

Anderson, L. R. and C. A. Holt (1997): "Information cascades in the laboratory," *The American economic review*, 847–862.

Arechar, A. A., S. Gächter, and L. Molleman (2018): "Conducting interactive experiments online," *Experimental Economics*, 21, 99–131.

Azrieli, Y., C. P. Chambers, and P. J. Healy (2018): "Incentives in experiments: A theoretical analysis," *Journal of Political Economy*, 126, 1472–1503.

Banerjee, A. V. (1992): "A simple model of herd behavior," *The quarterly journal of economics*, 107, 797–817.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of political Economy*, 100, 992–1026.

Caplin, A., D. Csaba, J. Leahy, and O. Nov (2018): "Rational inattention, competitive supply, and psychometrics," Tech. rep., NBER, #25224.

Çelen, B. and S. Kariv (2004): "Distinguishing informational cascades from herd behavior in the laboratory," *American Economic Review*, 94, 484–498.

Chandrasekhar, A. G., H. Larreguy, and J. P. Xandri (2018): "Testing models of social learning on networks: Evidence from Two Experiments," Tech. rep., Working Paper.

——— (2019): "Testing models of social learning on networks: Evidence from two experiments," *Econometrica*, forthcoming.

Charness, G., R. Oprea, and S. Yuksel (2018): "How do people choose between biased information sources? Evidence from a laboratory experiment," mimeo.

Chen, D. L., M. Schonger, and C. Wickens (2016): "oTreeAn open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, 9, 88–97.

Choi, S., D. Gale, and S. Kariv (2012): "Social learning in networks: a quantal response equilibrium analysis of experimental data," *Review of Economic Design*, 16, 135–157.

Cipriani, M. and A. Guarino (2005): "Herd behavior in a laboratory financial market," *American Economic Review*, 95, 1427–1443.

Coleman, J., E. Katz, and H. Menzel (1957): "The diffusion of an innovation among physicians," *Sociometry*, 20, 253–270.

Conley, T. G. and C. R. Udry (2010): "Learning about a new technology: Pineapple in Ghana," *American economic review*, 100, 35–69.

Drehmann, M., J. Oechssler, and A. Roider (2005): "Herding and contrarian behavior in financial markets: An internet experiment," *American Economic Review*, 95, 1403–1426.

——— (2007): "Herding with and without payoff externalitiesan internet experiment," *International Journal of Industrial Organization*, 25, 391–415.

Edwards, W. (1968): "Conservatism in human information processing," in *Formal Representation of Human Judgement*, ed. by B. Kleinmuntz, New York, 17–52.

Enke, B. and F. Zimmermann (2017): "Correlation neglect in belief formation," *The Review of Economic Studies*, 86, 313–332.

Eyster, E., M. Rabin, and G. Weizsäcker (2018): "An experiment on social mislearning," mimeo.

Eyster, E. and G. Weizsäcker (2016): "Correlation neglect in portfolio choice: Lab evidence," mimeo.

Gale, D. and S. Kariv (2003): "Bayesian learning in social networks," *Games and Economic Behavior*, 45, 329–346.

Goeree, J. K., T. R. Palfrey, B. W. Rogers, and R. D. McKelvey (2007): "Self-correcting information cascades," *The Review of Economic Studies*, 74, 733–762.

Grether, D. M. (1980): "Bayes rule as a descriptive model: The representativeness heuristic," *The Quarterly Journal of Economics*, 95, 537–557.

——— (1992): "Testing Bayes rule and the representativeness heuristic: Some experimental evidence," *Journal of Economic Behavior & Organization*, 17, 31–57.

Hara, K., A. Adams, K. Milland, S. Savage, C. Callison-Burch, and J. P. Bigham (2018): "A data-driven analysis of workers' earnings on amazon mechanical turk," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ACM, 449.

Harel, M., E. Mossel, P. Strack, and O. Tamuz (2018): "Rational Groupthink," mimeo.

Hossain, T. and R. Okui (2013): "The binarized scoring rule," *Review of Economic Studies*, 80, 984–1001.

Hung, A. and J. Dominitz (2004): "Homogeneous Actions and Hetergeneous Beliefs: Experimental Evidence on the Formation of Information Cascades," Tech. rep., Carnegie Mellon University.

Hung, A. A. and C. R. Plott (2001): "Information cascades: Replication and an extension to majority rule and conformity-rewarding institutions," *American Economic Review*, 91, 1508–1520.

Kübler, D. and G. Weizsäcker (2004): "Limited depth of reasoning and failure of cascade formation in the laboratory," *The Review of Economic Studies*, 71, 425–441.

Mueller-Frank, M. and C. Neri (2013): "Social Learning in Networks: Theory and Experiments," mimeo.

Snowberg, E. and L. Yariv (2018): "Testing the Waters: Behavior across Participant Pools," Tech. rep.

Tubbs, R. M., G. J. Gaeth, I. P. Levin, and L. A. Van Osdol (1993): "Order effects in belief updating with consistent and inconsistent evidence," *Journal of Behavioral Decision Making*, 6, 257–269.

A Implementation

The subject pool consists of U.S. workers on Amazon Mechanical Turk (MTurk) that previously completed at least 100 tasks with an overall approval rate of at least 95%.[34] Each subject is paid only if she completes the experiment.[35] The experiment comprises two parts. In the first part, subjects have up to 15 minutes to read the experimental instructions after which they have up to 8 minutes to complete an incentivized quiz, with each correct answer receiving $0.15.[36] Before accepting the task, subjects are told an approximate time (10 minutes) needed to complete the second part of the experiment. Subjects are also told that the decisions of their matched partner will not affect their chances to obtain the high bonuses. No other information about the second part of the experiment is provided until a subject decides to participate and is presented with the instructions. Each subject can only participate once.

After completing the quiz, subjects receive feedback on their answers.[37] If a subject decides to proceed with the experiment, she waits to be matched with another MTurk worker.[38] When two subjects are successfully matched, they are prompted to the second part of the experiment. In each of the 20 periods of the game, each subject has up to two minutes to submit her decisions for both the matching and guessing tasks. If a subject times out, the subject and her partner are terminated from the experiment and only the subject that did not trigger the timeout is paid.[39]

Given the positive probability that a no match might occur when subjects randomly arrive over time, subjects are required to wait for at least five minutes for a match to arrive. If no match occurs after five minutes, subjects are allowed to quit the study or wait for longer in the hope

[34] Conducting online experiments on MTurk has become more widespread in recent years. Arechar et al. (2018) compare the results of a public goods experiment conducted both in the laboratory as well as using subjects from MTurk and find that "basic patterns of behavior online are similar to those in the laboratory." Snowberg and Yariv (2018) conducted a larger scale experiment involving three different samples: Caltech undergraduates, U.S. MTurk workers, and a representative sample of the U.S. Population. They find that the behavior of MTurk workers is roughly in between that of Caltech undergraduates and the representative sample in terms of first-order stochastic dominance of the response distributions.

[35] If a subject's partner drops out, the subject is paid for one randomly chosen decision from each task until her partner dropped out. If a subject cannot be matched, she receives the minimum guaranteed payment of $1 plus her earnings from the quiz. The instructions clearly state all these contingencies.

[36] The instructions and the quiz can be found in the online appendix.

[37] The quiz is designed to direct a subject's attention to the most important components of the experiment, especially what a subject and her partner observe in every period of the experiment depending on the treatment.

[38] Matching subjects *after* completing the quiz is done to ensure that dropouts do not prevent their (committed) partners from being matched with other committed participants. A previous (but unrelated) online experiment revealed that dropping out is most likely after subjects read the instructions.

[39] Subjects are informed about this possibility from the start. The hard time constraint prevented subjects matched with a dropout from being stuck on a waiting page without the ability to submit the assignment and receive payment.

of getting matched. If a subject decides to quit, she is paid her participation fee (\$0.05),[40] her earnings from the quiz and the minimum guaranteed payment of \$1 to compensate the subject for her time.

B Omitted Tables and Figures

	Min	Median	Mean	Max	St. Dev.	# obs
All	0	5	4.51	6	1.37	492
Male	0	5	4.55	6	1.31	273
Female	0	5	4.55	6	1.43	220
Other	0	5	4.00	4	2.65	3

Table 5: Summary statistics of quiz scores.

[40]While admittedly low, the participation fee is in line with other studies using Amazon Mechanical Turk (e.g., Caplin et al, 2018).

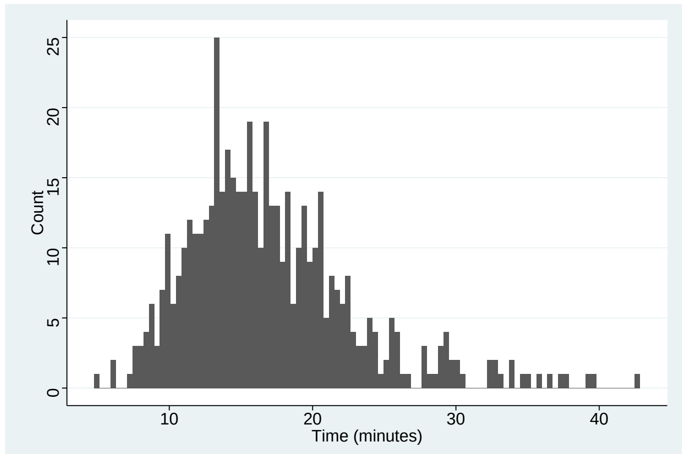


Figure 6: Distribution of time spent on the experiment at the subject-level.

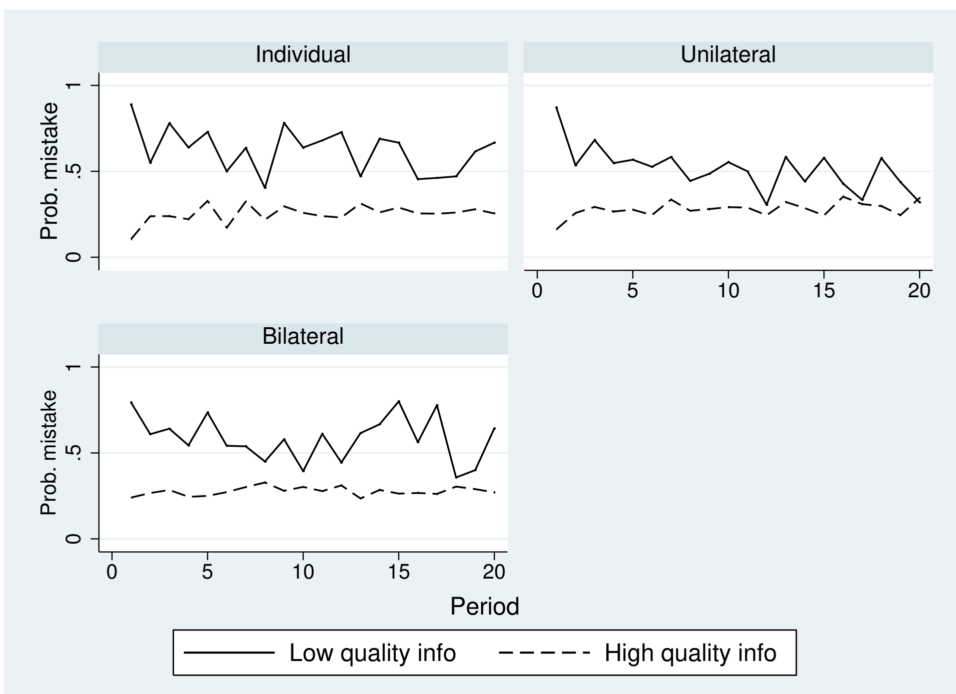


Figure 7: Mistakes by treatment based on quality of own information.

C Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
		Probit			IV Probit	
	Individual	Unilateral	Bilateral	Individual	Unilateral	Bilateral
Other player's action in $t-1$	0.0815	**0.315******	**0.214******	-0.00334	**0.544*****	**0.443****
	(0.0534)	(0.0645)	(0.0604)	(0.242)	(0.205)	(0.208)
Fraction of orange balls in t	0.822****	0.213	0.768****	0.903***	-0.113	0.251
	(0.228)	(0.200)	(0.187)	(0.311)	(0.268)	(0.258)
Orange ball in t	0.756****	0.821****	0.724****	0.769****	0.849****	0.822****
	(0.127)	(0.0960)	(0.107)	(0.136)	(0.105)	(0.115)
Fraction of orange actions in $t-1$	1.880****	1.658****	1.552****	2.525****	2.403****	2.276****
	(0.136)	(0.141)	(0.135)	(0.182)	(0.216)	(0.215)
Constant	-1.803****	-1.483****	-1.551****	-2.114****	-1.814****	-1.814****
	(0.115)	(0.115)	(0.109)	(0.140)	(0.132)	(0.125)
Observations	3173	3173	3078	2505	2505	2430

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 6: The effect of other players' actions (coefficients reported from probit and IV probit regressions).

	(1)	(2)	(3)	(4)	(5)	(6)
		OLS			IV	
	Individual	Unilateral	Bilateral	Individual	Unilateral	Bilateral
Other player's action in $t-1$	0.0156	**0.0740******	**0.0648*****	0.00452	**0.181*****	**0.128****
	(0.0172)	(0.0211)	(0.0212)	(0.0658)	(0.0643)	(0.0601)
Fraction of orange balls in t	0.0768	-0.00204	-0.111	0.0806	-0.0506	-0.139
	(0.113)	(0.117)	(0.105)	(0.114)	(0.121)	(0.109)
Orange ball in t	0.236****	0.276****	0.273****	0.237****	0.275****	0.273****
	(0.0408)	(0.0348)	(0.0371)	(0.0404)	(0.0345)	(0.0369)
Orange ball in $t-1$	0.0322	-0.00900	0.0159	0.0322	-0.00896	0.0165
	(0.0242)	(0.0233)	(0.0242)	(0.0242)	(0.0231)	(0.0241)
Orange ball in $t-2$	0.0184	0.00757	0.0431**	0.0187	0.00823	0.0413**
	(0.0218)	(0.0218)	(0.0203)	(0.0220)	(0.0217)	(0.0197)
Orange ball in $t-3$	0.0265	0.00960	0.0786****	0.0268	0.0102	0.0778****
	(0.0212)	(0.0184)	(0.0218)	(0.0213)	(0.0180)	(0.0217)
Orange ball in $t-4$	0.0351*	0.0332	0.0314	0.0355*	0.0355*	0.0304
	(0.0193)	(0.0210)	(0.0205)	(0.0194)	(0.0210)	(0.0203)
Orange ball in $t-5$	0.0113	-0.0251	0.0320*	0.0115	-0.0250	0.0319*
	(0.0191)	(0.0212)	(0.0189)	(0.0192)	(0.0214)	(0.0188)
Fraction of orange actions in $t-1$	0.747****	0.756****	0.707****	0.749****	0.722****	0.691****
	(0.0431)	(0.0495)	(0.0483)	(0.0446)	(0.0512)	(0.0503)
Constant	-0.106****	-0.0532*	-0.0403	-0.104****	-0.0664**	-0.0511*
	(0.0243)	(0.0290)	(0.0263)	(0.0261)	(0.0310)	(0.0289)
Observations	2505	2505	2430	2505	2505	2430

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 7: The effect of other players' actions (robustness checks with lagged balls).

	(1)	(2)	(3)	(4)	(5)	(6)
		OLS			IV	
	Individual	Unilateral	Bilateral	Individual	Unilateral	Bilateral
Other player's action in $t-1$	0.0153	**0.0690*** **	**0.0605*** **	-0.0171	**0.168** **	**0.120** **
	(0.0190)	(0.0230)	(0.0212)	(0.0757)	(0.0697)	(0.0601)
Fraction of orange balls in t	0.129	-0.100	0.0442	0.150	-0.148**	0.0145
	(0.0823)	(0.0626)	(0.0645)	(0.0964)	(0.0679)	(0.0648)
Orange ball in t	0.227****	0.284****	0.261****	0.227****	0.282****	0.260****
	(0.0417)	(0.0367)	(0.0383)	(0.0414)	(0.0368)	(0.0381)
Fraction of orange actions in $t-1$	0.227*	0.170	0.157	0.232*	0.140	0.154
	(0.127)	(0.142)	(0.130)	(0.125)	(0.147)	(0.128)
Orange action in $t-1$	0.144****	0.160****	0.148****	0.143****	0.159****	0.144****
	(0.0340)	(0.0343)	(0.0335)	(0.0338)	(0.0341)	(0.0327)
Orange action in $t-2$	0.0450	0.0791**	0.0561*	0.0465	0.0763**	0.0522*
	(0.0323)	(0.0313)	(0.0308)	(0.0325)	(0.0315)	(0.0299)
Orange action in $t-3$	0.118****	0.0916***	0.0870***	0.118****	0.0912***	0.0850***
	(0.0307)	(0.0306)	(0.0324)	(0.0305)	(0.0308)	(0.0321)
Orange action in $t-4$	0.103****	0.0890***	0.0978****	0.103****	0.0853**	0.0946***
	(0.0295)	(0.0340)	(0.0287)	(0.0295)	(0.0343)	(0.0290)
Orange action in $t-5$	0.0494	0.0411	0.0830**	0.0500*	0.0398	0.0845***
	(0.0299)	(0.0272)	(0.0320)	(0.0298)	(0.0272)	(0.0320)
Orange action in $t-6$	0.0567*	0.0527**	0.102***	0.0556*	0.0477*	0.101***
	(0.0323)	(0.0260)	(0.0338)	(0.0321)	(0.0263)	(0.0334)
Orange action in $t-7$	0.0370	0.0538**	0.0282	0.0372	0.0550**	0.0288
	(0.0345)	(0.0263)	(0.0302)	(0.0343)	(0.0261)	(0.0298)
Orange action in $t-8$	-0.00176	-0.000564	-0.000876	-0.000963	0.00384	-0.00344
	(0.0301)	(0.0306)	(0.0282)	(0.0304)	(0.0300)	(0.0280)
Orange action in $t-9$	-0.00409	0.0246	-0.0245	-0.00454	0.0252	-0.0236
	(0.0280)	(0.0236)	(0.0300)	(0.0279)	(0.0234)	(0.0300)
Orange action in $t-10$	-0.00438	0.0374	-0.00291	-0.00586	0.0434*	-0.00244
	(0.0286)	(0.0230)	(0.0232)	(0.0282)	(0.0239)	(0.0231)
Constant	-0.0724****	-0.0170	-0.0330	-0.0699****	-0.0253	-0.0417*
	(0.0200)	(0.0244)	(0.0200)	(0.0212)	(0.0274)	(0.0221)
Observations	1670	1670	1620	1670	1670	1620

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 8: The effect of other players' actions (robustness checks with lagged actions).

	(1)	(2)	(3)	(4)	(5)	(6)
		OLS			IV	
	Individual	Unilateral	Bilateral	Individual	Unilateral	Bilateral
Other player's action in $t-1$	0.00876	0.0844****	0.0481***	-0.0116	0.151**	0.136**
	(0.0161)	(0.0208)	(0.0184)	(0.0685)	(0.0643)	(0.0590)
Other player's action in $t-2$	0.0226	0.0122	0.0311	0.0530	0.0493	-0.00112
	(0.0179)	(0.0180)	(0.0213)	(0.0599)	(0.0565)	(0.0645)
Fraction of orange balls in t	0.249***	0.0564	0.217***	0.223**	-0.0321	0.0936
	(0.0766)	(0.0694)	(0.0678)	(0.0940)	(0.0890)	(0.0795)
Orange ball in t	0.230****	0.270****	0.236****	0.221****	0.272****	0.254****
	(0.0397)	(0.0324)	(0.0362)	(0.0400)	(0.0339)	(0.0362)
Fraction of orange actions in $t-1$	0.636****	0.623****	0.570****	0.722****	0.713****	0.667****
	(0.0399)	(0.0436)	(0.0440)	(0.0450)	(0.0527)	(0.0547)
Constant	-0.0797***	-0.0166	-0.0291	-0.110****	-0.0686**	-0.0482
	(0.0241)	(0.0288)	(0.0256)	(0.0274)	(0.0319)	(0.0309)
Observations	3006	3006	2916	2505	2505	2430

Subject-clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Table 9: The effect of other players' actions.

D Construction of the optimal action in the unilateral social learning treatment

Let $\Theta = \{\theta, \neg\theta\}$ be the binary state space and $T_i = T = \{t_\theta, t_{\neg\theta}\}$ be the binary signal space, and take an arbitrary period number n. Let p_θ denote the common prior belief that the state is θ. As in our experimental design, the signal generating technology is given by

$$\text{Prob}(t = t_\theta|\theta) = q_\theta > \frac{1}{2} > \text{Prob}(t = t_\theta|\neg\theta) = q_{\neg\theta}. \tag{3}$$

Suppose that a Bayesian agent (Player 1 from now on) in the unilateral social learning treatment observed n_θ balls of color θ up to (and including) period n as well as observing her partner's previous action (Player 2 from now on) a^2_{n-1}, with the partner also assumed to be Bayesian and Player 1 knowing that Player 2 is Bayesian.[41] Since we are assuming that both agents are Bayesian expected utility maximizers, a sufficient statistic for an arbitrary private history $\mathbf{t}^n$ is the number of signals of type t_θ out of n total signals observed, which we denote by n_θ. From now on, we refer to n_θ instead of $\mathbf{t}^n$.

[41] Past actions are also observable.

Let $p_1^n(\theta|n_\theta)$ denote Player 1's posterior belief that the state is θ after having observed n_θ signals of type t_θ. By Bayes' rule, this probability equals

$$p_1^n(\theta|n_\theta) = \frac{q_\theta^{n_\theta}(1-q_\theta)^{n-n_\theta}p_\theta}{q_\theta^{n_\theta}(1-q_\theta)^{n-n_\theta}p_\theta + q_{\neg\theta}^{n_\theta}(1-q_{\neg\theta})^{n-n_\theta}p_{\neg\theta}}. \tag{4}$$

Player 2 forms beliefs in a similar way.

We start by defining some notation. Let $r_{n-1}(n'_\theta|\theta)$ denote the probability of observing n'_θ signals of type t_θ out of $n-1$ signals, conditional on state θ, which equals

$$r_{n-1}(n'_\theta|\theta) = \binom{n-1}{n'_\theta} q_\theta^{n'_\theta}(1-q_\theta)^{n-1-n'_\theta}, \tag{5}$$

Next, suppose that Player 1 observes Player 2 having chosen action a_{n-1}^2 in period $n-1$. Since Player 2 cannot observe Player 1's past actions, her decision can only be based on her history of private signals. Let n'_θ denote the number of t_θ signals privately observed by Player 2. If $a_{n-1}^2 = \theta$, then it must be the case that $n'_\theta \geq \lceil \frac{n-1}{2} \rceil$. Let $E_\theta = \left\{ n'_\theta \geq \lceil \frac{n-1}{2} \rceil \right\}$ denote such an event. By Bayes' rule,

$$\text{Prob}_n\left(\theta|E_\theta\right) = \frac{\text{Prob}_n\left(E_\theta|\theta\right) p_1^n(\theta|n_\theta)}{\text{Prob}_n\left(E_\theta|\theta\right) p_1^n(\theta|n_\theta) + \text{Prob}_n\left(E_\theta|\neg\theta\right)\left(1 - p_1^n(\theta|n_\theta)\right)} \tag{6}$$

where

$$\text{Prob}_n\left(E_\theta|\theta'\right) = \sum_{n'_\theta = \lceil \frac{n-1}{2} \rceil}^{n-1} \text{Prob}(n'_\theta = k|\theta') = \sum_{n'_\theta = \lceil \frac{n-1}{2} \rceil}^{n-1} r_{n-1}(n'_\theta|\theta'), \quad \theta' \in \{\theta, \neg\theta\} \tag{7}$$

Similarly, if $a_{n-1}^2 = \neg\theta$, then

$$\text{Prob}_n\left(E_{\neg\theta}|\theta'\right) = \sum_{n'_\theta = 0}^{\lfloor \frac{n-1}{2} \rfloor} r_{n-1}(n'_\theta|\theta'), \quad \theta' \in \{\theta, \neg\theta\} \tag{8}$$