An Experiment on Social Learning with Information Sequencing*

Pëllumb Reshidi †

November 2023

Abstract

We test whether the order and timing of information arrival affect beliefs formed within a group. In a lab experiment, participants estimate a parameter of interest using a common and a private signal and past guesses of group members. Contrary to Bayesian predictions, participants strongly react to information sequencing, even though the informational content is unchanged. Though non-Bayesian, behavior is robustly predictable by a model relying on simple heuristics. We explore how observability of others' choices and information timing reduce correlation neglect. Lastly, we document a key heuristic—the influence of private information on participants' actions is time-independent.

*I thank Leeat Yariv, Roland Bénabou, Alessandro Lizzeri, Pietro Ortoleva, Guillaume Fréchette, Mark Dean, Jacopo Perego, Marina Agranov, Michael Thaler, Joao Thereze, and participants at many seminars and conferences for helpful comments and suggestions.

†Department of Economics, Duke University pellumb.reshidi@duke.edu

1 Introduction

When deciding which politician to vote for, which car to purchase, or which new restaurant to visit, people often turn to their social network for guidance. These networks often consist of various individuals who receive and share information at different times. When a group of individuals forms opinions on a topic of interest, do their conclusions remain consistent whether they initially receive the bad news, followed by intermediate news, and then the good news, in contrast to the reverse order? Do they also arrive at the same conclusion when receiving information sequentially rather than simultaneously? Does the network structure, namely, how socially connected group members are with one another, interact with the sequencing of information? These are the central questions explored in this paper.

We provide an experimental examination of how beliefs evolve within a group when information arrives sequentially. While keeping the objective information content unchanged, we study whether the timing and order in which this information arrives affect final beliefs.

In theory, when group members aiming to estimate the same parameter have ample time to deliberate and update their beliefs via Bayes' rule, the sequence of information arrival should have no impact on their final beliefs—all that should matter is the content of the information. However, whether this is the case in practice is unclear. Understanding if and to what extent the sequencing alone matters is important for numerous setups: a board of directors deciding whether to expand the presence of the corporation, facing sequential arrival of information due to different departments reporting at different times; a committee evaluating a job market candidate, incorporating information sequentially based on conference interactions, communication with advisors, CV, and research evaluations; a community evaluating a newly formed political party, with individuals acquiring information about the party members, and discussing with others throughout. In addition, a better grasp of how the sequencing of information matters can help us judge the robustness or fragility of information acquisition via decentralized social platforms—an increasing and already substantial source of information for many.[1]

At odds with Bayesian learning, when agents learn from their social environment through different heuristics, the particular sequencing of information may matter. Reshidi (2023) extends the classic DeGroot (1974) social-learning model to allow for sequential information arrival and illustrates that in such environments, the timing and order in which information arrives have a first-order effect on the group's final beliefs. In the current paper, we further build on the sequential version of the DeGroot model by allowing agents to adjust the weight

[1]In a survey with more than 9,200 respondents, Shearer and Mitchell (2021) find 71% of Americans get at least part of their news through social media platforms.

they place on other agents who recently received information.[2] Based on this augmented model, by changing the timing and order of signals while keeping the information content unchanged, we construct sequences predicted to sway the group towards forming upward-biased and downward-biased final beliefs. We then utilize these sequences to evaluate whether groups are affected by the sequencing of information and, if so, whether they are influenced in the direction predicted by the model.

Experimental Design. Each participant plays eight games, with each game consisting of 40 rounds. At the beginning of each game, groups of four participants observe a common signal informative of a payoff-relevant parameter. In each round, participants are incentivized to input their best guess of this parameter. Within a group, participants have access to a subset of other members' previous-period guesses. In the first treatment, participants can see the previous guess of only one other member (a ring network); in the second treatment, participants have access to the previous guesses of all members (a complete network). In addition to the common signal, each participant eventually receives a single informative private signal. These signals arrive in different sequences: in some games, all participants receive their signals jointly in the same round; in other games, information arrives sequentially, with each participant receiving a signal at a different round. Between each round of information arrival, there are multiple rounds in which information can disseminate through the network.

We vary whether the private signals arrive jointly or sequentially, and in the latter case, we also vary their arrival order. The game is designed such that all other elements remain fixed—the network, the realized signals, and the identity of those receiving them. In doing so, we isolate the effect that the timing and order of information per se have on the final beliefs formed by the group.

Results. First, the data shows that the order and timing of information arrival influence groups' final beliefs. The same group of participants, within the same network structure, receiving the same information in a different order or timing, form beliefs that are statistically significantly distinct. Therefore, the sequencing of information appears to have a substantial impact in social learning settings. For the specific context, this is at odds with Bayesian learning, under which we ought to see no difference.

Second, we find the ordering of information and the network structure can be utilized to reduce correlation neglect, a well-documented and undesirable phenomenon. We find that an increase in the connectivity of the network structure, from a ring to a complete network, can

[2] A relevant extension as the setup in the experiment utilizes a small network with explicit knowledge of the timing of information arrival.

reduce correlation neglect by up to 32%. This shift might, in practice, be a demanding one; such an increase in network connectivity might come with high additional costs. However, we find that releasing information sequentially instead of simultaneously, a potentially much more attainable change, can reduce correlation neglect by up to 75%.

Third, although deviations from optimal behavior manifest in various aspects, the predominant behavioral component driving the documented data seems to be a lack of participants' ability to adjust how strongly they incorporate their private signal depending on *when* they receive this signal. We document participants place excessive weights on their own signals, a phenomenon that has been documented in the literature in different ways and under different labels such as overconfidence, overreaction, and so on. However, we also find that participants do not adjust these weights over time. We believe to be the first to document this time-independent manner of incorporating one's own information. In doing so, we point out an important heuristic.

Finally, while the above findings are not tied to a specific model specification, we proceed to estimate a hybrid model that nests the Bayesian and sequential DeGroot models, while allowing for intermediate levels of sophistication. Our results indicate that the version of the sequential DeGroot model that allows for more flexible weights on others' previous actions does an impressive job of predicting participants' behavior. This suggests that information sequencing influences beliefs in a predictable manner. While additional testing and model refinements may be warranted, the robustness of the current model's predictions is nevertheless striking.

Literature. A considerable experimental literature (in the lab and in the field) studies social-learning stylized facts and which social-learning models best describe the observed data. Papers such as Chandrasekhar et al. (2020), Choi et al. (2005), Eyster et al. (2015), Mobius et al. (2015), Mueller-Frank and Neri (2013), and Grimm and Mengel (2020) conduct experiments comparing naive with Bayesian social-learning models. In these papers, all information arrives jointly at the beginning of the experiment. Typically, as long as the network structure becomes modestly complex or the information regarding the network structure is limited, the Bayesian learning model no longer predicts the participants' behavior well. Agranov et al. (2021) study a social-learning setup in which each participant in each round receives a new informative binary signal. They vary whether participants can see others' guesses and signals, as well as the group size and network structure. They find both group size and observability of others' actions and signals improve final outcomes.

Of relevance is also the literature on correlation neglect. Kallir and Sonsino (2009), Ortoleva and Snowberg (2015), Levy and Razin (2015), and Enke and Zimmermann (2017) emphasize the prevalence of correlation neglect. As Enke and Zimmermann (2017) show, correlation neglect is present even in simple environments where information sources are mechanically correlated. Dasaratha and He (2019) analyze the extent to which correlation neglect affects social learning in a vertical network structure, in which participants can observe previous players' actions and take a single action. They find that a naive social-learning model does a better job of predicting various comparative statics.

Most existing experimental literature employs discrete signals and actions (typically binary).[3] In such cases, because the action space is coarse, participants' actions cannot fully convey their beliefs. Using a discrete action space, especially if the choices are binary, makes it difficult, if not impossible, to point estimating weights participants place on their own signals or on other participants' actions. Beyond the repeated-actions setup, the rich action space is an additional key difference from the herding literature, in which, in the typical framework, because the action space is coarse, the exact beliefs cannot be conveyed, leading to an eventual information cascade. However, some papers do utilize a continuous-action space, such as Corazzini et al. (2012) and Brandts et al. (2015).[4] By varying the network structure, these papers study whether predictions originating from naive social-learning models hold, such as that more central individuals have greater influence.

What sets apart our experiment from this work is the sequential arrival of information, as opposed to all information being delivered in the first round.

Of relevance is also the literature on individual-level belief updating. For an excellent summary, see Benjamin (2019). Out of the many documented errors in probabilistic reasoning and judgment biases, the most relevant to our work are errors exhibited by individuals who update their beliefs after receiving multiple signals sequentially, in particular, their tendencies to overweight earlier or later signals. Several experiments suggest that people are more influenced by information that they receive early in a sequence compared to information that they receive in the middle of the sequence. This is known as the *primacy effect*, see Peterson and DuCharme (1967), Roby (1967), Dale (1968), and De Swart and Tonkens (1977). On the other hand, several studies observe a *recency effect*, suggesting that the most recent signals have a greater influence on final beliefs than those observed in the middle of a sequence, see Pitz and Reinhold (1968), Shanteau (1970), Marks and Clarkson (1972), Edenborough (1975), and Grether (1992).

An important distinction from this literature is that in our setting, participants receive a single private signal. Thus, from an individual updating point of view, there is no room

[3]In some instances, the signal space is rich, drawn from, say, a normal distribution, but actions are still discrete, such as guessing left or right.

[4]Angrisani et al. (2018) also employ a continuous action-space in an experiment with sequential learning, in which, agents act one after the other and move only once.

for a *primacy* or *recency effect*. However, in our setup, the time when the private signal is received differs, allowing us to measure whether participants adequately adjust the weight they place on the only private signal they receive. Since participants receive a single signal, this weight adjustment is thus present only in social learning settings, and can not directly be tested in a single-agent setup.

To the best of our knowledge, this paper is the first to take the timing of information arrival in a social learning environment seriously and to set up an experiment in which the sequencing of information can be altered while keeping all other components unchanged.

2 Conceptual Framework

This section introduces the environment, Bayesian agents' behavior, and relevant non-Bayesian models. A reader not interested in the predictive power of non-Bayesian models can skip their discussion. All model-free results, such as whether the timing and order of information arrival affect final beliefs, changes in correlation neglect due to changes in the network structure or information sequencing, and how participants are affected by their private signals, can be interpreted as particular deviations from the Bayesian benchmark. For readers interested in the non-Bayesian models, especially in the estimation of the parameterized model in section 5, the rest of this section is relevant.

The Environment. Consider a directed network of four agents, where a link from agent i to agent j implies agent i can observe the past action of agent j. In particular, consider a *ring network* and a *complete network*. At the beginning of the game, a payoff-relevant parameter θ is drawn from a uniform distribution on the interval $[0, 1000]$. The game consists of 40 rounds. In the first round, all agents receive a common signal s_c, normally distributed with mean θ and variance σ_θ^2. In addition, agents receive private signals s_i, which are also normally distributed with mean θ and variance σ_i^2. To simplify the environment, we set $\sigma_i^2 = \sigma_\theta^2 = \sigma^2$, making each signal equally informative. Unlike the common signal, which arrives in the first round, the private signals arrive in rounds 5, 13, 21, or 29. We refer to these rounds of the game as the *information rounds*. Signals arrive either sequentially, each signal in a distinct information round, or simultaneously: all signals arriving jointly in round 5. After 40 rounds, for one randomly chosen round $\tilde{r}$, with each round being equally likely to be chosen, the payoff is calculated as

$$\text{payoff}_i = \max\left\{\, B - \gamma\left|\theta - g_{i,\tilde{r}}\right|, 0 \,\right\},$$

where $\gamma > 0$ represents the sensitivity of the payoff with respect to the distance from the θ target. B represents the maximum bonus, which is achieved if the guess $g_{i,\tilde{r}}$, of agent i in the randomly chosen round $\tilde{r}$ is exactly equal to θ. The payoff linearly decreases from B the further the guess is from θ, with a lower bound of zero.

Bayesian Agents. Consider a group of Bayesian agents who have common knowledge of Bayesian rationality. With only the common signal available, regardless of their risk preferences, Bayesian agents maximize their expected payoff by setting their guess $g_{i,r} = s_c$. To see this, note that because the prior is diffused, the distribution of the common signal represents the posterior distribution of the beliefs of the agents. Hence, after observing common signal s_c, Bayesian agents believe θ is distributed normally with mean s_c and variance σ^2.[5] The expected payoff is then the dot product of the symmetric and single-peaked posterior with the symmetric and single-peaked payoff function. Furthermore, notice the payoff function remains single-peaked and symmetric regardless of whether the agent is risk-neutral, risk-averse, or even risk-seeking. Because the dot product is supermodular, the expected payment is maximized by arranging the posterior beliefs with the payoff function, which happens when the guess $g_{i,r}$ is set equal to the mean of the posterior beliefs. For a formal proof of this result, see section 7.1 in the Appendix.

When information arrives sequentially, the first agent receiving a private signal s_i will form beliefs that are normally distributed with mean $\frac{s_c+s_i}{2}$ and variance $\frac{\sigma^2}{2}$. Via the same argument as above, the optimal guess of this agent is $g_{i,r} = \frac{s_c+s_i}{2}$. In the complete network, with a one-period delay, all other agents observe the new guess of agent i. By inverting the guess, they are able to learn the signal s_i perfectly and, thus, also incorporate it in their guess. In the ring network, agent j, the immediate neighbor of agent i, with a one-period delay, observes the new guess of agent i. Once more, agent j would be able to invert the guess, perfectly learn s_i, and incorporate it in their new guess. The immediate neighbor of agent j would then follow the same procedure, so would their neighbor, and so on, until the signal s_i is incorporated by all. For both the complete and the ring network, the same procedure follows when the second, third, and fourth signals are released.

Similarly, when information arrives simultaneously, initially, all agents optimally incorporate their private information. Afterward, in the complete network, agents see the new guesses of all other agents, invert these guesses, and incorporate the signals of all agents.

[5] Although the uniform distribution is not a conjugate prior of the normal distribution, for realizations of θ sufficiently away from the boundaries, the posterior of Bayesian agents is approximately normally distributed with mean s_c and variance σ^2. In what follows, we focus on realizations of θ that are sufficiently away from the 0 and 1000 boundaries, making the described behavior a good approximation. We discuss why these are the relevant cases in the experiment-design section. For a numerical approximation of the optimal guess given a signal, see the Online Appendix.

In the ring network, after incorporating their own signal, the agent inverts the guess of the immediate neighbor and incorporates their signal in the next round. In the next round, by observing the change in their neighbor's guess, they invert the signal their neighbor's neighbor must have received. They do so for another round until the signals of all agents are included in the guess.

Notice that after each round of information release, be it of a single signal or all signals jointly, the new information is incorporated by all agents within two(four) rounds in the complete(ring) network.[6] After n private signals are released, once the information is incorporated by the agents, the variance of the posterior shrinks to $\frac{\sigma^2}{1+n}$, while the guess is re-optimized to the average of the n signals. Thus, in both networks, after all, information is released and disseminated, the guesses of all agents converge to $\frac{s_c+\sum_{i=1}^{4} s_i}{5}$—regardless of the order of signal release, or whether the signals were released simultaneously or sequentially.

Sequential DeGroot Agents (SD Agents). As in the classic DeGroot (1974) social-learning model, consider agents who, in each period, form new guesses by taking a convex combination of their own and their observable neighbors' previous-period guesses. Let $m_{i,j}$ represent the weight agent i places on agent j, with $m_{i,i}$ representing the weight agent i places on their own previous-period guess. If agent i cannot observe a particular agent j directly, the weight $m_{i,j}$ will be equal to 0. Weights are normalized to sum up to 1, $\sum_{j=1}^{N} m_{i,j} = 1$. To allow for sequential information arrival, assume that once agent i's private signal arrives, they place weight λ_i on their signal. Let $\gamma(t)$ represent the set of agents for whom signals arrive in round t, which will be an empty set whenever $t \notin \{5, 13, 21, 29\}$. The guess of agent i in round t is then

$$g_{i,t} = \begin{cases} \sum_{j=1}^{N} m_{i,j}\, g_{j,t-1} & if \quad i \neq \gamma(t) \\ (1-\lambda_i)\left(\sum_{j=1}^{N} m_{i,j}\, g_{j,t-1}\right) + \lambda_i\, s_i & if \quad i \in \gamma(t) \end{cases} . \tag{1}$$

Reshidi (2023) studies this setup. Under such learning dynamics, the main takeaway of relevance for this paper is that the final consensus formed by the agents depends on the timing and order of information arrival. That is, keeping the realized signals unchanged but switching the order of signal arrival, or whether information arrives sequentially instead of simultaneously, affects the final beliefs formed within the group.[7] This finding is in sharp contrast to the Bayesian predictions.

Although the DeGroot model has found wide applicability, both in theoretical work and

[6] Since there are eight rounds of updating in between each information release round, Bayesian agents have ample time to reach a consensus.

[7] For a detailed analysis of the model, see Reshidi (2023).

in practice, the model is limited by a key assumption, which is that the weights agents place on their neighbors, how much they *listen to them*, remains fixed over time; this heuristic might be reasonable when agents have incomplete information of the network structure or of the timing of information arrival. However, in the experiment analyzed in this paper, the network size is small and commonly known by all participants; moreover, participants are explicitly informed of the information-arrival rounds. Thus, it is expected that participants might change how much they pay attention to others based on whether others received information. To accommodate these changes, we extend the above model to allow for such weight shifts.

Reactive Sequential DeGroot Agents (RSD Agents). These agents act similarly to SD agents, except that they may modify the weights placed on previous-period guesses depending on the timing of information arrival. Let $\tilde{\gamma}(t)$ represent the set of agents who receive new information, be it their own private signal or observe a neighbor who received information in the previous round. The weight agent i assigns to agent j is

$$m_{i,j}(t) = \frac{\hat{m}_{i,j}(t)}{\sum_j \hat{m}_{i,j}(t)}, \qquad \hat{m}_{i,j}(t) = \begin{cases} \underline{m}_{i,j} & if \quad j \notin \tilde{\gamma}(t) \\ \alpha\, \underline{m}_{i,j} & if \quad j \in \tilde{\gamma}(t) \end{cases}. \tag{2}$$

When $\alpha > 1$ agents *pay more attention* to agents who recently received information. The guess of agent i in round t is then

$$g_{i,t} = \begin{cases} \sum_{j=1}^{N} m_{i,j}(t)\, g_{j,t-1} & if \quad i \neq \gamma(t) \\ (1-\lambda_i)\left(\sum_{j=1}^{N} m_{i,j}(t)\, g_{j,t-1}\right) + \lambda_i\, s_i & if \quad i \in \gamma(t) \end{cases}, \tag{3}$$

where the difference between the current specification (3) and the previous specification (1) is that the weights $m_{i,t}(t)$ are potentially time dependent, as described by equation (2). The relevant aspects of the behavior of such agents are analyzed in section 7.2 in the Appendix. Given the experiment features discussed above, small network size, and common knowledge of information-arrival, this specification is the one that is hypothesized, based on which the information sequencing is designed.

Hybrid Agents. Finally, we introduce hybrid agents that nest the Bayesian, SD, as well as RSD agents. For particular parameter values, the hybrid agents act as any of the above. We introduce these agents with the purpose of estimating a general model allowing the estimated parameter values to reveal which model best describes the behavior observed in the lab.

As was the case for RSD agents, the hybrid agents' specification also allows for different weights on group members who recently received information. In addition, this specification allows agents to place different weights on their signal given the signal arrival round; that is, the estimated $\lambda_{i,\hat{t}}$ is allowed to vary in time for the four information rounds $\hat{t} \in \{5, 13, 21, 29\}$. Furthermore, agents are allowed to anchor toward the common signal as well as their own signal.[8] Let δ_c represent the anchoring value toward the common signal, and let δ_s represent the anchoring value towards the private signal once it arrives. If complete convergence is not observed in the data, these parameters will allow us to see in which direction the guesses of the participants differ.

Let $\hat{t}_i$ represent the round in which agent i receives their private signal. The guess of agent i in round t is then

$$g_{i,t} = \begin{cases} (1-\delta_c)\left(\sum_{j=1}^{N} m_{i,j}(t)\, g_{j,t-1}\right) + \delta_c\, s_c & if \quad t < \hat{t}_i \\ \left(1-\lambda_{i,\hat{t}}\right)\left((1-\delta_c)\left(\sum_{j=1}^{N} m_{i,j}(t)\, g_{j,t-1}\right) + \delta_c\, s_c\right) + \lambda_{i,\hat{t}}\, s_i & if \quad t = \hat{t}_i \\ (1-\delta_c-\delta_s)\left(\sum_{j=1}^{N} m_{i,j}(t)\, g_{j,t-1}\right) + \delta_c\, s_c + \delta_s s_i & if \quad t > \hat{t}_i \end{cases}, \tag{4}$$

where $m_{i,j}(t)$ is once more determined by (2). We revisit the hybrid model in section 5, where we estimate the aforementioned parameters and see which model best describes the behavior of the participants in the lab.

3 Experiment Design

The experiment was coded on *Python* using *oTree* from Chen et al. (2016), an open-source platform for experimental design. A description of the interface and sample instructions are available in the Online Appendix. The experiment design received approval from Princeton University IRB.

The Truth and Signals. Participants play a total of eight games, each consisting of 40 rounds. They are informed that θ, referred to as *the truth*, is drawn from a uniform distribution on $[0, 1000]$.[9] Participants are not told the realized value of θ. They receive

[8] Friedkin and Johnsen (1990) consider a model in which, although the beliefs of the agents become more and more similar, they do not necessarily converge. This lack of convergence happens if agents place a permanent weight on their own signal. Consequently, although a common component of actions converges, all actions are *anchored* toward the private signals.

[9] Because the realized value of θ plays no role for identification, its values are drawn once and are equal to $\{455, 793, 312, 126, 202, 871, 312, 644, 542\}$.

a common signal s_c as well as a private signal s_i.[10] Participants are informed that these signals are normally distributed around θ with a standard deviation of 30. The common signal is observed by all group members, whereas the private signal is observed only by the participant who receives it. Thus, within each game, each group receives five different but equally informative signals.

The Timing of Signals. In the first round of each game, the common signal is displayed to all participants. Unlike the common signal, private signals arrive in information rounds 5, 13, 21, or 29. The purpose of the seven rounds between each information round is to ensure ample time is available for information to disseminate across the group.[11] The exception is the gap between the common signal and the first signal (rounds 1 and 5), which is only three rounds. Because the common signal is the same for all participants, allowing ample time for information to disseminate is unnecessary because all participants share the same information, namely, the common signal.[12]

In six of the eight games, information arrives sequentially, with each participant receiving their private signal in one of the information rounds. In two of the eight games, all signals arrive jointly in round 5.

Guesses, Incentives, and Feedback. Participants are asked to input their best guess of θ within each round. To incentivize them to report their beliefs truthfully, we implemented the following payment scheme: within *three* randomly chosen games, *one* out of 40 rounds is randomly chosen (with each game and round being equally likely to be chosen), and payoffs are calculated as follows:

$$\text{payoff} = \max\left\{ \$\,10 - \frac{1}{4}\left|\theta_j - g_{i,j,r}\right|, 0 \right\},$$

where θ_j represents the truth, or the target value, in game j, and $g_{i,j,r}$ represents the participant's guess in round r of game j. The payoff linearly decreases as the guess is further away from the truth, with a lower payoff bound of \$0. Regardless of risk preferences, the expected payoff of the participant is maximized by reporting the mean of the posterior beliefs. See section 7.1 in the Appendix for the optimal guess proof.

[10]Instead of having a prior that is normally distributed, we choose to have a diffused prior with a common signal that is normally distributed. We do so to minimize Base-rate neglect. The effectiveness of this approach in reducing Base-rate neglect is studied in Agranov and Reshidi (2023).

[11]Thus, including the round in which the signal arrives, there are four times(two times) more rounds than needed for a group of Bayesian agents to reach a consensus in the complete(ring) network.

[12]However, we do not shrink this gap further because we want to ensure the common signal is exclusively displayed to all participants as long as any other signal.

At the end of each game, participants are informed of the true value of θ, but not whether the game that just ended is among the three randomly chosen games that will be used for payment.

Grouping Procedure. Participants are informed that they will be paired with new group members after a game ends. In the first four games $j \in \{1, 2, 3, 4\}$, groups of four participants are formed randomly within the session; however, in the last four games $j \in \{5, 6, 7, 8\}$, participants are regrouped in the same groups as in the first four games. That is, for $j > 4$, participants are grouped in the same groups in round j as they were in round $j - 4$. The motives for this particular grouping protocol will be explained shortly.

Signal Generation. In the first four games $j \in \{1, 2, 3, 4\}$, at the beginning of the game, five signals are drawn from a normal distribution with mean θ and variance 30.[13] Each signal is afterward assigned to be either the common signal or the signal of a particular participant. For $j > 4$, let $c_j = \theta_j - \theta_{j-4}$ represent the difference between the realized target value in round j from the realized target value in round $j - 4$. In game $j > 4$, we reuse the realized signals from game $j - 4$ and shift them by c_j so that they are centered around the new target value. That is,

$$c_j = \theta_j - \theta_{j-4} \qquad s_{c,j} = s_{c,j-4} + c_j \qquad s_{i,j} = s_{i,j-4} + c_j \quad i \in \{1, 2, 3, 4\},$$

with $s_{c,j}$ representing the common signal and $s_{i,j}$ the private signals of participant i in round j. Shifting the signals by the common c_j constant ensures the signals remain normally distributed around the new target value θ_j. However, the spread of the signals is the same as it was four rounds earlier when the same participants were in the same group: the difference between a participant's signal from the common signal or from their neighbors' signals remains unchanged.[14] Consider an example of possible realized signals in rounds 3 and 7:

$$\begin{array}{llllll} s_{c,3} = 318 & s_{1,3} = 270 & s_{2,3} = 302 & s_{3,3} = 350 & s_{4,3} = 334 & \theta_3 = 312, \\ s_{c,7} = 650 & s_{1,7} = 602 & s_{2,7} = 634 & s_{3,7} = 682 & s_{4,7} = 666 & \theta_7 = 644. \end{array}$$

Note subtracting $332 (c_7 = \theta_7 - \theta_3)$ from signals in round 7, and θ_7 leads to the signals received in round 3 and θ_3. Thus, adding this constant after the data collection process is equivalent

[13] The expected range for five normally distributed signals with standard deviation 30 is approximately 70. To ensure a relatively representative draw, if the range of the five signals is lower than 56 or higher than 84, a new sample was drawn.

[14] Because participants only see the common signal and their private signal, identifying the repetition of the spread of signals is not possible.

to the group receiving the exact same signals. The motives for recycling the signals in this way will be explained shortly.

Signal Arrival Order. Even with four participants and only four information release rounds, conditional on the realized signals, 4^4 sequences of information release are possible. Although for Bayesian agents, all these sequences would lead to the same final consensus, even for non-Bayesian agents, numerous such sequences would affect beliefs in a similar way. Therefore, to properly test whether beliefs are being updated in a non-Bayesian way, we focus on sequences where, at least theoretically, we anticipate disparities in the final beliefs.[15]

As emphasized in section 2, given the features of the experiment (small network size, common knowledge of the network's structure, and common knowledge of the timing of information arrival), it is to be expected that participants might change how much they pay attention to other group members based on information arrival. Therefore, the conjectured behavior is that of RSD agents. The predictions for RSD agents are analyzed in section 7.2 in the Appendix. For reasonable parameter values, the key takeaways of relevance are that: (i)releasing a signal in a later round increases its influence on the final beliefs, and (ii)releasing signals jointly increases the influence the common signal has on the final beliefs. Motivated by these predictions, we construct the following sequences.

SEQ_U - (Sequential, Up) - Signals arrive sequentially, with the lowest realized signal arriving in round 5, followed by the second lowest in round 13, second highest in round 21, and highest in round 29.

SEQ_D - (Sequential, Down) - Signals arrive sequentially, with the highest realized signal arriving in round 5, followed by the second highest in round 13, second lowest in round 21, and lowest in round 29.

SIM_U - (Simultaneous, Up) - All participants receive their signals in round 5, with the highest value signal being assigned to the common signal.

SIM_D - (Simultaneous, Down) - All participants receive their signals in round 5, with the lowest value signal being assigned to the common signal.

Based on the comparison presented in section 7.2.3 in the Appendix, sequences denoted by U, for up, are expected to generate final beliefs that converge to a higher level than sequences denoted by D, for down. These comparisons follow from the two predictions described above. If signals released in later rounds influence the final beliefs more, releasing the same information via SEQ_U is expected to generate higher final beliefs than releasing this information via SEQ_D. On the other hand, if a joint release of signals leads to excessive

[15]Note that if we were to randomly draw sequences, by drawing sequences that lead to the same final beliefs, we would often wrongly conclude that participants are updating beliefs in a Bayesian manner even if they were not doing so.

influence of the common signal on the final beliefs, when $s_c > \max_i s_i$, a joint release SIM_U is expected to generate higher final beliefs than a sequential release via SEQ_D. Similar logic justifies the comparison between SIM_D and SEQ_U.[16]

In each session, in three games, information is released as SEQ_U, in three games, information is released as SEQ_D, information in one game is released as SIM_U, and in one game information is released as SIM_D. In each session, the first four games alternate between *up sequences* and *down sequences*. The choice of the last four sequences is closely linked to the first four sequences. If the sequence for game $j \in \{1, 2, 3, 4\}$ was an up sequence, the sequence for game $j + 4$ is a down sequence, and vice versa. An example is presented in figure 1.[17]

Figure 1: Sequence Matching

As a result, two of the matched sequences are *SEQ vs SEQ*, that is, sequential up versus sequential down sequences, and two of the matched sequences are *SIM vs SEQ*, consisting of a simultaneous up matched with a sequential down sequence and a simultaneous down matched with a sequential up sequence. In total, there are four matched sequences corresponding to eight games.

Putting Everything Together. In game $j \in \{5, 6, 7, 8\}$, groups are identical to those in game $j - 4$. This regrouping ensures the difference in the group's final guesses in game j and $j - 4$ does not reflect differences in individual characteristics. For each group in game $j \in \{5, 6, 7, 8\}$, the truth θ and the signals are identical to those in game $j - 4$.[18] The recycling of signals ensures the difference in the group's final guesses in game j and $j - 4$ is not a result of the particular signal realizations. Because the network structure is fixed within a session, the network structure between games $j \in \{5, 6, 7, 8\}$ and $j - 4$ is also unchanged. Keeping the network structure unchanged within a session ensures the difference in the group's final guesses in games j and $j - 4$ is not a result of the varying network structure.

Then, within each group, the only difference in game $j \in \{5, 6, 7, 8\}$ and game $j - 4$ is

[16] For a detailed analysis, see section 7.2 in the Appendix.

[17] To ensure the effect we capture is not driven by the order of up and down sequences, the order of the sequences differs in different sessions. The alternative utilized sequence starts with a down sequence: $\{SEQ_D, SEQ_U, SEQ_D, SIM_U, SEQ_U, SIM_D, SEQ_U, SEQ_D\}$.

[18] Except for a common shifter c_j, which is subtracted after the data collection.

the sequencing of information arrival.[19] Therefore, we have created a controlled environment that allows us to isolate the effect that information sequencing has on the group's final beliefs.

Nonetheless, a difference between game $j \in \{5, 6, 7, 8\}$ and $j - 4$ is the participant's accumulated experience. Participants may have understood the game better and thus modified their strategies. This concern is what motivates the alternating sequences. The distribution of up and down sequences across the first and last four rounds is equal; thus, participants' experience cannot selectively affect the final beliefs of upward or downward sequences. For any reasonable impact that learning might have on participants' strategies, this alternation between up and down sequences should ensure learning is not what drives the results.

Although the experimental setup might seem somewhat complicated from our point of view as analysts, the setup remains rather simple from the participant's perspective. In essence, participants are asked to guess a parameter value. They are provided with signals to aid in their guesses and can also see the guesses made by other group members.

Treatments and the Interface. In theory, it is sufficient to display the signal for a single round for participants to incorporate it; nonetheless, to ensure participants indeed observe and incorporate the signals, both the common and the private signals, once they arrive, are displayed for four consecutive rounds. The interface allows agents to see the last two guesses of their observable group members. Had we allowed participants to observe only the previous-period guesses of their neighbors, Bayesian agents would need at least one period recall to make optimal choices. However, by providing the last two past guesses of their neighbors, Bayesian agents can make optimal decisions within each round with zero recall of the actions taken in the previous periods. In other words, all the relevant information needed for an optimal choice at any time is available on the screen. For a detailed breakdown of the interface, see the Online Appendix.

To account for environments in which the available information content differs, we run two treatments, the *low-information treatment* and the *high-information treatment*. We do so as a robustness check. That is, does the phenomenon we document prevail only in conditions with limited information, only in conditions with abundant information, or both?

In the low-information treatment, the network structure is a ring network: participants see the past guesses of only one group member, whereas their past guesses are observed by another group member. Participants have no uncertainty about the timing of the common signal; it always arrives in the first round. They also have no uncertainty about whether they have received their signal. However, in rounds in which a signal arrives, but the participant

[19] The shuffling of the order of the sequences (whether the session starts with an upward or a downward sequence) ensures the documented difference is not driven by the order itself.

themselves is not the signal recipient, they are not explicitly informed about the identity of the group member who received the signal. In theory, by observing their neighbor's guess changes, they can deduce whether the neighbor they observe is the one who just received a signal.[20]

In the high-information treatment, the network structure is a complete network: participants see past guesses of all other group members, whereas their past guesses are observed by all. Once more, participants know when the common and their private signal arrives. In addition, in rounds in which a signal arrives, but the participant themselves is not the signal recipient, they are explicitly informed about the identity of the group member who received the signal. Thus, in the high-information treatment, participants have more information through more observable past guesses, as well as through the knowledge of the identity of the signal recipient.

Experiment Summary. Experiments were run at *Princeton Experimental Laboratory for the Social Sciences* (PExL). In total, 136 students participated in the experiment, with 64 participants in the low-information treatment and 72 participants in the high-information treatment. We ran at least four sessions for each treatment, with at least 12 participants in each session. The average earnings were \$31, including the \$10 show-up fee.[21]

4 Results

4.1 Bayesian Expected Behavior

Before we explore the results from the observed data, we simulate the behavior we would see from Bayesian players (as described in section 2).

Recall that participants play eight games, leading to four matches of upward versus downward sequences. The average optimal Bayesian guesses are shown in figure 2. To make graphing these values easier, we normalize the signals and guesses by subtracting the group's common signal from them.[22] Thus, in each graph, zero corresponds to the group's common signal, which is why, in the beginning, before any other signal arrives, the average group guess is equal to zero. The gray, dash-connected dots represent the normalized average

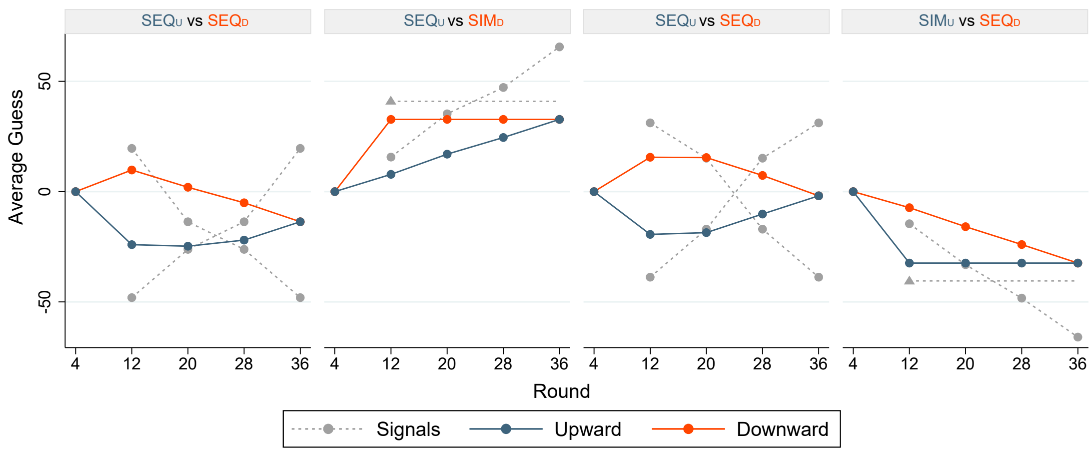
[20] Furthermore, they know information arrives sequentially in six out of eight games, whereas in two out of eight games, all information arrives jointly in round 5. Thus, if a participant does not receive a signal in round 5, they can deduce that they are not in a round in which all signals arrive jointly in that round.

[21] The highest theoretical earning is \$40, achieved if the participant exactly guessed θ on the randomly chosen rounds in all three randomly chosen games.

[22] To ensure an adequate comparison with figures we introduce later on, we use realized signal values from the actual experiment.

signals for sequences that release information sequentially. The gray triangle and the dashed line represent the average of all private signals for sequences that release information jointly. The blue(orange) connected dots represent the normalized average beliefs of Bayesian agents after the signals have been received via an upward(downward) sequence. Recall that signals arrive in rounds 5, 13, 21, or 29. To allow for enough time for information to disseminate in the network, the average group beliefs have been calculated seven rounds after any signal release. The guesses are averaged across all rounds with the same sequence matching.

Figure 2: Average Bayesian Guesses



Within each of the four matched sequences, the information content is always identical; the only difference is the order in which information arrives. As figure 2 shows, regardless of whether signals arrive via upward sequences (blue graphs) or via downward sequences (orange graphs) after all information has been received and disseminated in the network, the final beliefs of the group converge to the same value. In other words, the final beliefs formed by Bayesian players are not affected by the order of information arrival, nor are they affected by whether signals arrive sequentially or simultaneously.

In the online appendix, we simulate the behavior of Bayesian agents that might make mistakes (implementation errors). We do so by adding noise on the signal incorporation weights (λ weights), as well as on the weights placed on others' observable actions ($m_{i,j}$ weights). Regardless of the amount of implementation noise, on average, the final beliefs from upward and downward sequences converge to the same point. Thus, the addition of noise, while making the data less precise, does not change the predicted Bayesian behavior.

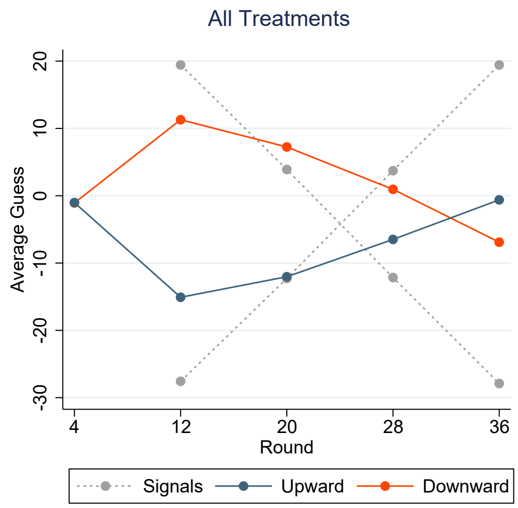
Note, however, that even for Bayesian agents, the beliefs formed from upward and downward sequences differ before all information has been released. This difference should come as no surprise because, until that point, participants see different information. Not until *all* information has been released does the information content of upward and downward sequence become identical.

It is also important to note that the evolution of beliefs is not necessarily monotonic even in cases in which information is released sequentially in an increasing or decreasing manner, see for example the second and third graph in Figure 2. This is because the common signal lies between the extreme signal values.

4.2 Information Sequencing and Final Beliefs

Cross Plots. We now explore how participants' actual beliefs, or reported guesses, are affected by the sequencing of information. Figure 3 plots participants' guesses averaged over all matched sequences across both treatments. It is similar to figure 2; however, it aggregates all matched sequences and utilizes the actual data instead of the simulated Bayesian guesses.

Figure 3: Observed Average Guesses across Treatments



If the sequence of signal arrival played no role in determining final beliefs, both graphs should converge to the same value, as was the case for the simulated Bayesian guesses. However, from figure 3, we see that the graphs do not converge to the same value, implying that the sequencing affected final beliefs. Thus, even though the network structure, the group members, and the received signals were identical, the order in which these signals were received influenced the final beliefs formed within the group.

Furthermore, the figure reveals beliefs are affected in a predictable manner: the upward sequences (blue graph), on average, converge to higher beliefs than the downward sequences (orange graph). Recall that these sequences were a key part of the experimental design. They were based on the predicted behavior of RSD agents and constructed before any data was collected. The above graph reveals that, on average, over all matched sequences, across both treatments, the RSD model seems to do a good job predicting the direction in which information sequencing affects final beliefs. Furthermore, figure 6 in section 7.3.1 in the Appendix shows the predicted direction holds separately both in the low- and high-information treatments. Even more striking, the predicted direction seems to hold, on average, not only within treatment but also within each matched sequence in each treatment. This breakdown is presented in figure 7 in section 7.3.1.

Although we anticipated that the constructed sequences would (i)affect final beliefs and (ii)be able to somewhat predict the direction of influence, we did not expect that the predicted direction would prove to be correct on average within each matched sequence. This finding is striking, and we interpret it to mean participants' behavior, although not Bayesian, is robustly predictable.

Sequence Influence Regression. Although the above graphs show that, on average, the upward sequences converge to higher values than the downward sequences, from these graphs alone, we cannot discern if this difference is statistically significant. To see if a statistically significant difference exists between the beliefs generated by the different sequences, we subtract the downward sequence from the matched upward sequence. Note there are seven rounds between the arrival of private signals. However, 11 rounds take place between the arrival of the last signal and the end of the game (rounds 29 and 40). To exclude the possibility that the last signal has a higher impact because of this longer dissemination period, the main analysis will utilize data from rounds 1 to 36. This approach ensures equal rounds of information dissemination after each private signal release.[23]

Although not the focus of the main analysis, we analyze the evolution of beliefs through each round in section 7.3.2 in the Appendix. There we further examine the difference between the evolution of beliefs in the complete and ring network.

Recall that the plots above show that after all information is released and sufficient time has passed for information to disseminate in the network, the beliefs originating from the upward sequences converge to higher values than those originating from the downward sequences. For the high-information treatment, information can fully disseminate one additional round after the last signal is released, whereas, for the low-information treatment, information can fully disseminate within three additional rounds after the last signal is released. Thus, any round from 32 onward can be used to test this hypothesis. For the reasons specified above, we focus on rounds 32-36. In Table 1, we present a fixed-effects regression, allowing for different intercepts in each round.[24] The variable *Constant* captures the difference between the beliefs from upward and downward sequences. *Sequential* is a dummy variable equal to 1 if the matched sequence was *SEQ vs SEQ*, that is, if both the matched

[23]The additional rounds towards the end of the game are utilized in the analysis in the appendix, in which we show that the captured differences are not temporary.

[24]In the Online Appendix, the last four rounds are analyzed to ensure that no major difference is observed.

sequences released information sequentially. *Low Info* is a dummy variable equal to 1 if the data comes from the low-information treatment, whereas *Sequential* × *Low Info* represents an interaction term between these two variables. Each column presents the same regression with a different error clustering level.[25]

Table 1: Sequence Influence

	$\triangle$ *Guess*: Round 32-36		
	(No Cluster)	(Individual Cluster)	(Group Cluster)
Constant	8.298***	8.298***	8.298***
	(0.000)	(0.000)	(0.000)
Sequential	0.151	0.151	0.151
	(0.855)	(0.899)	(0.941)
Low Info	-0.180	-0.180	-0.180
	(0.833)	(0.916)	(0.941)
Sequential × *Low Info*	-1.952	-1.952	-1.952
	(0.105)	(0.366)	(0.594)
Round Fixed Effects	Yes	Yes	Yes
N	2720	2720	2720

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The estimated difference is statistically significant, with a p-value much lower than 0.01 regardless of the clustering level. The estimated coefficients for *Sequential*, *Low Info*, and *Sequential*× *Low Info* are not statistically significant, implying the gap between the upward and downwards sequences is roughly the same in both treatments and matched sequences.[26] Depending on the round, the treatment, and the sequence, the gap is approximately between five and eight and always statistically significant. To put things into perspective and to evaluate whether this effect is economically significant, note the initial standard deviation of each signal is 30; because the group receives five signals, the ex-post standard deviation, or the standard deviation that the posterior should have given the information available to the group, is then $^{30}/\sqrt{5} \approx 13.4$. Hence, the estimated effect approximately corresponds to at least $^{1}/_{3}$ and at most $^{2}/_{3}$ of the ex-post standard deviation. We argue this effect is sizable, given that the network structure, the group members, and the received signals were identical, yet, by changing the order in which information is presented, we can have a first-order effect on the final beliefs.

4.3 Early and Delayed Signals

We next evaluate the intensity with which a signal affects final beliefs depending on when the signal was released. We utilize data from games in which information is released sequentially.

[25] To make sure that the significance documented in Table 1 is not being driven by pooling the observed data across different rounds, **Table 1** in the Online Appendix reports a regression utilizing only data in round 36—results remain qualitatively unchanged. Furthermore, as can be seen in figure 8, figure 9, and figure 10 show, the difference is significant in any of these rounds.

[26] Round 33 and 34 fixed effects are not statistically significant. Round 35 fixed effect is equal to -1.12 and is significant at the 0.10 level, whereas round 36 fixed effect is equal to -1.51 and significant at the 0.05 level.

We compare the weight signals have on participants' guesses after all information has been released and disseminated in the network. We allow for a level shift on the signals' estimated effect for games in which the signals were released in later rounds while also controlling for the common signal. The regression reported in Table 2 presents the estimated parameter values separately for the low and high-information treatments. Alternative p-values are also presented depending on the clustering level of the errors.[27]

Table 2: Earlier vs Later Signals

	Guess: : Round 32-36					
	low-information treatment			*high-information treatment*		
	(No C)	(Individual C)	(Group C)	(No C)	(Individual C)	(Group C)
S_c	0.282***	0.282***	0.282***	0.225***	0.225***	0.225***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
S_i	0.148***	0.148***	0.148***	0.152***	0.152***	0.152***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Additional Weight	0.0554***	0.0554***	0.0554**	0.0848***	0.0848***	0.0848***
	(0.000)	(0.001)	(0.017)	(0.000)	(0.000)	(0.000)
N	1920	1920	1920	2160	2160	2160

p-values in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The S_c parameter captures how the common signal affects final beliefs, S_i captures how a private signal affects final beliefs if released in an early round(round 1 or 2), and the *Additional Weight* variable captures the additional influence a signal has on the final beliefs if this signal is released in a later round(round 3 or 4). The positive and statistically significant value of the *Additional Weight* indicates that when released in later rounds, signals affect the final beliefs more. In particular, this weight boost is 0.0554 in the low-information treatment and 0.0848 in the high-information treatment. Thus, for example, if a signal is released in an earlier round, the signal's influence on the final beliefs is about 0.15. However, if the same signal is released in a later round, the weight of this same signal would be boosted to 0.20(0.23) in the low-(high-)information treatment. Hence, indeed, signals released in later rounds seem to affect final beliefs more.

Because the common signal and the private signals have the same precision, after information disseminates, the optimal weight to place on each signal is $1/5$. Thus, in the low-information treatment, the common signal is overweighted, the earlier released signals are underweighted, and the later released signals are, on average, adequately weighted. On the other hand, in the high-information treatment, the early released signals are underweighted once more, whereas both the common and later released signals are overweighted. In the following subsection, we further study the common signal's influence on final beliefs.

[27]To make sure that the significance documented in Table 2 is not driven by pooling the observed data across different rounds, in the Online Appendix, we report regressions utilizing only single round data; results remain qualitatively unchanged.

4.4 The Common Signal and Correlation Neglect

We further examine how the common signal influences the final beliefs formed within a group. To do so, we regress participants' guesses, after all information has been released and disseminated in the network, on the common signal S_c, and interactions of S_c with SEQ and HIT. SEQ is an indicator equal to 1 if the data originates from a game in which information was released sequentially. SIM, which stands for simultaneous, is an indicator function equal to $1 - SEQ$. HIT is an indicator equal to 1 if the data originates from the high-information treatment. We also control for private signals and their interaction with the above-mentioned indicators. Because the variables of interest are the common signal and its interactions with the above-mentioned indicators, these estimated values are presented in Table 3. The complete regression is reported in **Table 4** in the Online Appendix.

Table 3: Weight on Common Signal

	Guess: Round 32-36		
	(No Cluster)	(Individual Cluster)	(Group Cluster)
S_c	0.412***	0.412***	0.412***
	(0.000)	(0.000)	(0.000)
$S_c \times SIM \times HIT$	-0.0677***	-0.0677***	-0.0677***
	(0.000)	(0.025)	(0.159)
$S_c \times SEQ$	-0.159***	-0.159***	-0.159**
	(0.000)	(0.000)	(0.001)
$S_c \times SEQ \times HIT$	-0.0737**	-0.0737**	-0.0737*
	(0.000)	(0.006)	(0.032)
Private Signals Controls	Yes	Yes	Yes
N	5440	5440	5440

p-values in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

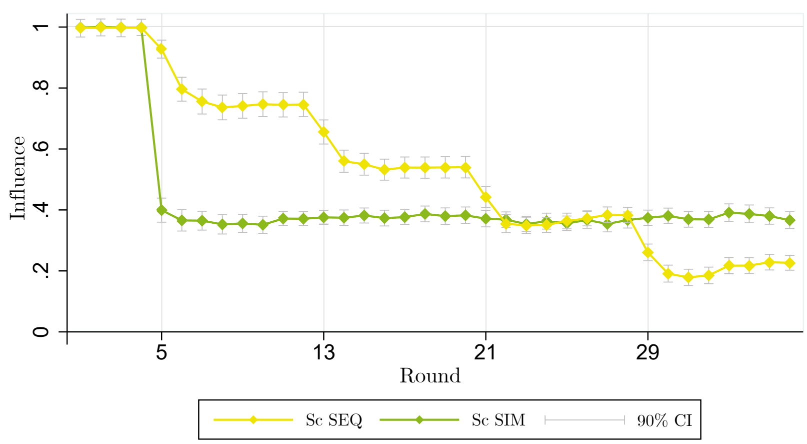
The coefficient on S_c captures the common signal's influence when information is released simultaneously in the low-information treatment. The common signal accounts for about 41.2% of the composition of the final beliefs. Recall that because the common and private signals have the same precision, the optimal weight to place on each signal is $1/5$. Thus, in the low-information treatment, when information arrives simultaneously, participants are excessively influenced by the common signal. Although participants learn from the common and private signals, they also rely on other group members' guesses when updating their beliefs. In our setup, correlation neglect, a well studied and documented phenomenon, arises if participants do not adequately account for the fact that others' guesses are correlated. This correlation arises from other group members also incorporating the common signal in their guesses.

As can be seen from the above regression, overweighting the common signal, or correlation neglect, is about 32% lower in the high-information treatment.[28] However, these two

[28]The estimated parameter on $S_c \times SIM \times HIT$ is -0.0677, whereas the correlation neglect, or the additional weight on S_c beyond it's optimal level, is $0.412 - 0.2 = 0.212$. This additional weight is reduced by $-0.0677/0.212 \approx 0.32\%$.

treatments may be thought of as separate setups, because fundamentally changing the network structure in practice, especially in the direction of making all nodes more connected, might not be a feasible policy change. Table 3 reveals that a sequential release of information can also drastically reduce correlation neglect. Concretely, when going from a simultaneous release of information to a sequential release of information, correlation neglect is reduced by about 75%.[29] In contrast to fundamentally changing the network structure, altering the order of information release, depending on the institution, might be a feasible policy change. By releasing information sequentially, the magnitude of correlation neglect may be sizably reduced.

Figure 4: The Common Signal's Influence



The graph above represents the weight the common signal has on participants' guesses in each round. The graph also shows the 95% confidence interval of the estimated weight.

Figure 4 presents how the common signal's influence evolves through each round. As can be seen, when information arrives sequentially, the weight commanded by the common signal gradually declines after each round in which a private signal arrives. By contrast, we see that in games in which all information arrives jointly, the dynamics of the common signal's influence are quite different. In such games, the common signal loses a substantial amount of influence immediately in round 5, when all participants receive their private signals. However, afterward, this impact barely changes, and consequently, the common signal ends up overinfluencing the final beliefs.[30] To summarize, by avoiding the release of information in large batches, especially in nodes that are directly connected, correlation neglect may be substantially reduced. Depending on the institution's network structure, this may be achieved without increasing the total information dissemination time.

[29] The estimated parameter on $S_c \times SEQ$ is -0.159, leading to a reduction of ${}^{-0.159}/_{0.212} \approx 0.75\%$.

[30] For a breakdown of the common signal's influence by treatment, see the Online Appendix.

4.5 Further Reduced Form Support for RSD Agents

Recall that the particular sequences utilized in this experiment were motivated by two predictions from the RSD agents model. For reasonable parameter values, the key takeaways of relevance are that: (i)releasing a signal in a later round increases its influence on the final beliefs, and (ii)releasing signals jointly increases the influence the common signal has on the final beliefs. As shown in Table 2 and in Table 3, both predictions hold in the data. Thus, the reactive sequential DeGroot model not only does a good job in predicting the direction in which information sequencing influences final beliefs, but it also does a good job in more fundamental predictions, such as how the timing of information affects a particular signal's influence. Importantly, the specific matched sequences utilized are locked in before any data is collected; therefore, we view this as a test of the model rather than a retrospective justification of the data.

5 Hybrid Model Estimation

5.1 Model Parameters

In this section, we estimate the hybrid-agents model described in section 2. The model nests the benchmark Bayesian agents, SD agents, and RSD agents while also allowing for intermediate levels of sophistication. Importantly, our intention is not to directly juxtapose the Hybrid model with the Bayesian model, the former, having more parameter values, naturally will outperform the latter. Instead, we can view this exercise as pitting the Bayesian model against the SD and RSD models. By examining the estimated parameter values, we can determine if private information is appropriately integrated, how information from others is utilized, and whether there are additional deviations from optimality, such as anchoring on one's own signal. We next describe the model parameters that we estimate.

Weights on Private Signals. In the estimated model, we allow participants to assign different weights to their private signal based on the round in which they receive their signal. The parameters to be estimated are $\{\lambda_5, \lambda_{13}, \lambda_{21}, \lambda_{25}\}$, where $\lambda_{\hat{t}}$ represents how much the signal influences the participants guess in the round it is received, if this signal is received in round $\hat{t} \in \{5, 13, 21, 29\}$.

Weights on Previous Guesses. Each round, participants input their best guesses and observe their own and other participants' previous guesses. In rounds in which no information arrives, we estimate the weight participants place on their own previous guess $m_{i,i}$ as well as the weight they place on other group members' previous guesses $m_{i,j}$. The low-information treatment utilizes a ring network; thus, there is only one $m_{i,j}$ with $j \neq i$ parameter to be estimated because participants see the past guesses of only one other group member. On the other hand, in the high-information treatment, participants see the past guesses of all other group members. However, because no reasonable distinction exists between the other group members beyond their guesses, we once more estimate one parameter $m_{i,j}$, which stands for the weight that participant i places on any other participant $j \neq i$.

In rounds in which others receive information, we estimate separate parameter values to evaluate whether participants change how much they pay attention to them. We estimate $\underline{m}_{i,i}$, $\underline{m}_{i,j}$, and $\overline{m}_{i,j}$, which are weights participants place on their own previous guess, on the guesses of group members who did not receive information last period, and on the guesses of group members who received information, last period, respectively.[31]

Anchoring parameters. Lastly, the hybrid model allows participants to anchor on the common and their own private signals. The parameters associated with such anchoring are δ_c and δ_s, respectively. Anchoring captures to what extent participants' guesses are directly affected by the common and their private signal (long after they have received these signals), beyond what can be explained as a convex combination of previous guesses.

5.2 Estimation Results

We present the estimated parameter values separately for the low- and high-information treatments in Table 4 and Table 5, respectively. For a clear separation of the estimated parameters, we use data from games in which information arrives sequentially. The Online Appendix goes through the estimation procedure in detail.

Table 4: Estimated Parameters: *Low-Information treatment*

	Weight on Private Signal			
	λ_5	λ_{13}	λ_{21}	λ_{29}
Value	0.60	0.61	0.59	0.54
95% *CI*	(0.53-0.68)	(0.49-0.73)	(0.49-0.70)	(0.46-0.62)

	Weight on Own Previous Guesses				Anchoring Parameters	
	$m_{i,i}$	$m_{i,j}$	$\underline{m}_{i,i}$	$\overline{m}_{i,j}$	δ_c	δ_s
Value	0.66	0.25	0.47	0.47	0.05	0.13
95% *CI*	(0.59-0.71)	(0.20-0.30)	(0.37-0.56)	(0.39-0.55)	(0.03-0.07)	(0.10-0.16)

Individual-level clustering

[31] Notice that in the round in which a participant receives information, they are the only ones that can react to it. Therefore, it will take at least one more round for other participants to react by seeing the new guess of the participant who received information. We account for this delay in the estimation procedure.

Table 5: Estimated Parameters: *High-information treatment*

	Weight on Private Signal			
	λ_5	λ_{13}	λ_{21}	λ_{29}
Value	0.70	0.57	0.64	0.61
95% *CI*	(0.61-0.80)	(0.50-0.64)	(0.54-0.75)	(0.53-0.70)

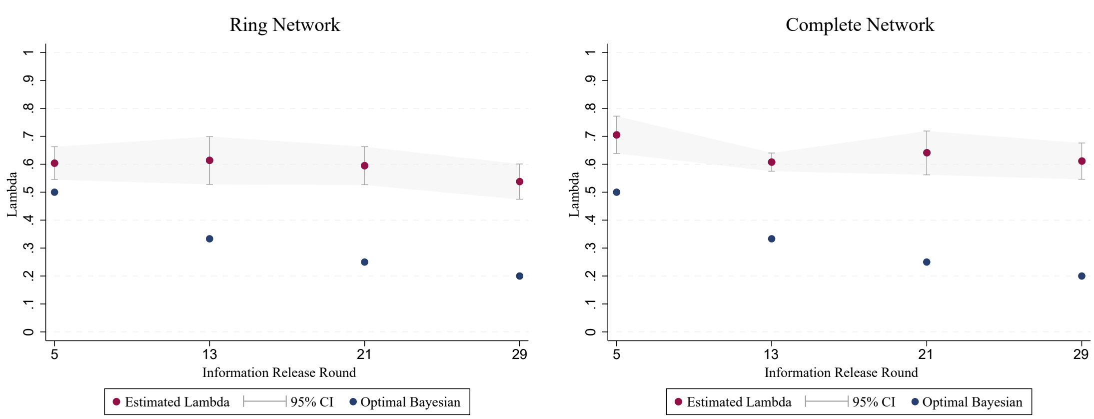
	Weight on Own Previous Guesses					Anchoring Parameters	
	$m_{i,i}$	$m_{i,j}$	$\underline{m}_{i,i}$	$\underline{m}_{i,j}$	$\overline{m}_{i,j}$	δ_c	δ_s
Value	0.57	0.13	0.31	0.07	0.50	0.02	0.07
95% *CI*	(0.52-0.61)	(0.11-0.14)	(0.22-0.40)	(0.03-0.11)	(0.46-0.54)	(0.01-0.03)	(0.06-0.08)

Individual-level clustering

Weights on Private Signals. Two stark features emerge regarding the weight participants place on their private signal. First, in both treatments, each estimated weight is greater than $1/2$. Recall that because the precision of the common signal is equal to the precision of the private signals, even if the participant is the first to receive a private signal, placing a weight greater than $1/2$ on the signal would not be optimal. Hence, we see a great deal of overreaction, or overconfidence, regarding participants' own information.

Second, going from λ_5 all the way to λ_{29}, the weights do not seem to sizably decrease. The first participant to receive a signal should be greatly influenced by it because the total information available is relatively low. However, participants receiving their signal in later rounds have incorporated others' information and thus should be influenced by their signal much less. For Bayesian agents these weights would be equal to $\{1/2, 1/3, 1/4, 1/5\}$, respectively. The roughly constant value of $\lambda_{\hat{t}}$ implies participants fail to comprehend that the weights they place on their signal should be a function of the arrival time, despite the fact that their signal's precision is unchanged. Figure 5 shows these weights graphically alongside the optimal Bayesian weights.

Figure 5: The Weight on the Private Signal



Although the literature has documented overreaction to one's own information, see Peterson and Miller (1965), Grether (1992), Ambuehl and Li (2018), we believe we are the first to document participants' inability to adjust their reaction based on the timing of information—a new and potentially important heuristic.[32]

Weights on Previous Guesses. Regarding the weights participants place on previous guesses, we see changes between the weights used in regular rounds and the weights used one round after a neighbor receives information. In the low-information treatment, participants increase their weight on the neighbor's guess from 0.25 to 0.47 in rounds when the neighbor recently received information. Consequently, in such rounds, participants decrease the weight they place on their own past guess from 0.66 to 0.47. Similarly, in the high-information treatment, participants increase from 0.13 to 0.50 the weight on a neighbor who received information. In such rounds, participants decrease the weight on their previous period guess from 0.57 to 0.31 while also decreasing from 0.13 to 0.07 the weight they place on other group members who did not receive information.[33]

Although far from optimal, in both treatments, participants react by increasing the weight on the guess of a group member who recently received information. Participants realize that when a group member receives new information, *paying more attention to them* is beneficial. Naturally, this change of weights could be viewed as a degree of sophistication on the participants' side.

Anchoring parameters. Finally, in both treatments, we estimate anchoring parameters that are statistically different from zero. The fact that the anchoring parameter on the participant's own signal δ_s is not equal to zero can explain why we do not see full convergence of participants' guesses, although the participants see and incorporate each others' previous guesses for a considerable number of rounds. We analyze belief convergence in detail in the Online Appendix.

In sum, the estimated model reveals that by paying more attention to group members who recently received information, participants show some degree of sophistication. In contrast,

[32] To reiterate from the literature review section, there is a literature on individual-level belief updating that finds a *primacy* and a *recency effect* when participants receive multiple private signals. In our setting, participants receive a single private signal. Thus, from an individual updating point of view, there is no room for a *primacy* or *recency effect*. However, in our setup, the time at which the private signal is received differs, allowing us to measure whether participants adequately adjust the weight they place on the only private signal they receive.

[33] Recall that in the high-information treatment, participants observe the past guesses of all other group members. Thus, the estimated $m_{i,j} = 0,13$ implies that in regular rounds, on average, participants place weight 0.13 on each of the other group members. On the other hand, in rounds in which a neighbor received information, the participant places weight 0.50 on that particular group member and 0.07 on the other two.

participants overreact and fail to adjust how much they are influenced by their own signal depending on the timing of the signal's arrival.[34]

6 Discussion

Although the experimental setup comprises multiple components and a somewhat complex design, from the participant's point of view, the experiment is rather simple. The small network size and the common knowledge of the timing of information arrival help maintain simplicity. Moreover, the common knowledge that all signals are equally informative greatly reduces the computational requirements for optimal behavior. Yet, despite the simple environment, we see sizable deviations from optimal behavior. Naturally, as is the case with every experiment, we wonder about the external validity of the results derived in this work. In practice, social networks are usually much larger than those utilized in this experiment. The network structure is typically not common knowledge, nor is the exact timing of when individuals receive information. Furthermore, the informativeness of the information sources across the network is typically not identical; some individuals have access to reliable information sources, others less so. This added complexity greatly increases the necessary computations for behavior to be anywhere near optimal. Thus, in practice, we expect even larger deviations from optimal behavior than those documented in this experiment. Consequently, the relevance of the sequencing of information increases with the size of these deviations. In other words, we believe our experiment captures a lower bound of the influence information sequencing has on beliefs formed through social learning.

Understanding how information sequencing influences social learning is important for numerous setups; besides those emphasized in the introduction, implications exist regarding the reliance on information acquisition via decentralized social platforms—an increasing and already substantial source of information for many.[35] Traditionally, news outlets were considered biased if they released a selected subset of the available information. However, in light of the importance of information sequencing, even news outlets releasing all available information may still have room to manipulate the social consensus simply by controlling the order and timing of information release. Given the current trends, keeping information sequencing in mind when designing and evaluating information platforms seems wise.

Our results point to several possible future directions of inquiry. Although we demon-

[34] These predictions are once more in line with the RSD model. In the Online Appendix, we analyze the best response to the data, that is, what the behavior of a Bayesian agent in a group composed of such participants would be. We find that while closer to, the behavior differs from the documented behavior.

[35] In a survey with more than 9,200 respondents, Shearer and Mitchell (2021) find 71% of Americans get at least part of their news through social media platforms.

strate the observed behavior is not in line with optimal behavior, we have shown it is nevertheless well predicted by a model relying on simple heuristics. Building on these initial steps, additional theoretical refinements of the model and supplementary empirical tests would be beneficial. Furthermore, seeing how the effects change with group size, other network structures, and potentially the inclusion of strategic elements, for example, an agent deciding whether they will view another agent's actions, would be of interest. This interaction between endogenous network formation and sequential information arrival might point to additional novel mechanisms. Naturally, a larger-scale field examination would be intriguing, for example, gathering data from social media platforms and studying how sentiment for a particular topic changed depending on the timing of positive, neutral, and negative news release.

7 Appendix

7.1 Proof of Optimal Guess

Once more, the payoff is

$$\text{payoff} = p(\theta, g) = \max\left\{ B - \gamma\left|\theta - g\right|, 0 \right\}.$$

We do not impose risk neutrality, the actual enjoyment of the payoff is some function $u\left(p(\theta, g)\right)$. The only condition we impose on $u(\cdot)$ is that utility is monotonic in payoff, earning more leads to higher utility. Without loss of generality, we normalize $u\left(0\right) = 0$. The realized θ value is unknown to the agent. Let $f(\theta)$ represent the probability distribution of θ. We assume that $f(\theta)$ is single-peaked and symmetric around this peak. Consequently, the peak will be equal both to the mean and median of the distribution, which we denote by μ. Naturally, this set of distributions includes, but is not limited to, the normal distribution. We show that it is optimal for an agent to guess $g^* = \mu$ by showing that any deviation from this guess leads to a lower expected utility. Consider a deviation $g > \mu$

$$\begin{aligned}
\mathbb{E}\left[u\left(p(\theta, \mu)\right)\right] - \mathbb{E}\left[u\left(p(\theta, g)\right)\right] &= \int u\left(p(\theta, \mu)\right) f(\theta) d\theta - \int u\left(p(\theta, g)\right) f(\theta) d\theta \\
&= \int \left(u\left(p(\theta, \mu)\right) - u\left(p(\theta, g)\right)\right) f(\theta) d\theta.
\end{aligned}$$

We know that $u\left(p(\theta, g)\right)$ is equal to zero for any $\theta \leq g - {}^{B}\!/_{\gamma}$, increasing as θ increases from $g - {}^{B}\!/_{\gamma}$ to g, decreasing as θ increases from g to $g + {}^{B}\!/_{\gamma}$, and is equal to zero for any value of $\theta \geq g + {}^{B}\!/_{\gamma}$. From the symmetry of the payoff function, and the assumption that utility is increasing in payoff, it follows that $u(p(\theta, g))$ is also single peaked and symmetric around g. The function $u(p(\theta, g))$ is the same function as $u(p(\theta, \mu))$ just shifted to the right by $g - \mu$. Let $d(\theta, \mu, g) = u(p(\theta, \mu)) - u(p(\theta, g))$. It then follows that $d(\theta, \mu, g)$ is nonnegative for $\theta \in \left(\mu - {}^{B}/_{\gamma} \frac{\mu+g}{2}\right)$, nonpositive for $\theta \in \left(\frac{\mu+g}{2}, g + {}^{B}/_{\gamma}\right)$ and zero everywhere else. The absolute value of this difference is symmetric around $\frac{\mu+g}{2}$.

$$
\begin{aligned}
\mathbb{E}\left[u\left(p(\theta, \mu)\right)\right] - \mathbb{E}\left[u\left(p(\theta, g)\right)\right] &= \int d(\theta, \mu, g) f(\theta) d\theta \\
&= \int_{\mu - B/\gamma}^{\frac{\mu+g}{2}} d(\theta, \mu, g) f(\theta) d\theta + \int_{\frac{\mu+g}{2}}^{g + B/\gamma} d(\theta, \mu, g) f(\theta) d\theta \\
&= \int_{\mu - B/\gamma}^{\frac{\mu+g}{2}} d(\theta, \mu, g) f(\theta) d\theta - \int_{\mu - B/\gamma}^{\frac{\mu+g}{2}} d(\theta, \mu, g) f\left(\theta - (\mu + g)\right) d\theta
\end{aligned}
$$

The equality of the first and second rows simply internalizes that $d(\theta, \mu, g)$ is equal to zero everywhere beyond the integration interval. Recall that $\frac{\mu+g}{2}$ splits the nonnegative region of $d(\theta, \mu, g)$ from it's nonpositive region. The equality of the second and third rows utilizes the fact that the absolute value of $d(\theta, \mu, g)$ is symmetric around $\frac{\mu+g}{2}$. Furthermore, the associated region of the probability distribution after changing the integration region of the *flipped* $d(\theta, \mu, g)$ becomes $f\left(\theta - (\mu + g)\right)$, which follows from the fact that $f(\theta)$ is symmetric and single peaked. Finally, we have

$$
\mathbb{E}\left[u\left(p(\theta, \mu)\right)\right] - \mathbb{E}\left[u\left(p(\theta, g)\right)\right] = \int_{\mu - B/\gamma}^{\frac{\mu+g}{2}} d(\theta, \mu, g) \left(f(\theta) - f\left(\theta - (\mu + g)\right)\right) d\theta > 0.
$$

The inequality follows from the fact that $d(\theta, \mu, g)$ is positive in the integrating region, while $f(\theta) - f\left(\theta - (\mu + g)\right)$ is positive for any value of $\theta < \frac{\mu+g}{2}$. It then follows that the integral must be positive. By symmetry, the result hold for deviations $g < \mu$.

□

7.2 Behavior of RSD Agents

7.2.1 Signal Weights in Sequential Information Release Rounds

Let signal i be released in *information round* i. From Reshidi (2023), we know that for SD agents, as long as the network matrix M is strongly connected and aperiodic, the final consensus will be

$$
c^{(K)} = \sum_{i=1}^{N} \left(\prod_{j=i+1}^{K} \left(1 - \pi_j \lambda_j\right) \right) \pi_i \lambda_i s_i + \prod_{i=1}^{N} \left(1 - \pi_i \lambda_i\right) c^{(0)}.
$$

Where $c^{(0)}$ in the current setting represents the common signal while s_i represents the signal released in information round i; π_i represents the social influence of the agent whose signal was released in information round i. More technically, π_i represents the i'th entry of the eigenvector corresponding to $\pi M = \pi$.

With RSD agents, for $z \in \{1, 2, \ldots, N\}$ rounds after a round in which a signal is released, the weights agents place on one another may differ from M, as these agents may place higher weights on their neighbors who just received information. The value of z depends on the network structure. In the complete network $z = 1$, while in the ring network $z = N - 1$.

Note that the weight changes depend on which agent received information. Let $\gamma(j)$ represent the agent who receives their signal in information round j. Let $M^{\gamma(j),t}$ represent the matrix of modified weights participants use $t \in \{1, z\}$ rounds after participant $\gamma(j)$ receives a signal. Let $M^{\gamma(j)} = M^{\infty} \prod_{t=1}^{z} M^{\gamma(j),t}$. Further let $c^{(j-1)}$ represent the existing consensus before signal $\gamma(j)$ was released. After signal $s_{\gamma(j)}$ is released, and participants communicate with one another, beliefs converge to

$$c^{(j)} = M^{\gamma(j)} \left(c^{(j-1)}, \ldots, \lambda_{\gamma(j)} s_{\gamma(j)} + (1 - \lambda_{\gamma(j)}) c^{(j-1)}, \ldots, c^{(j-1)}\right)'$$

Note that $M^{\gamma(j)}$ will be a matrix whose rows are identical and sum to one. The influence that signal $s_{\gamma(j)}$ will have on the new consensus is proportional to the $(\gamma(j), k)$'th entry of matrix $M^{\gamma(j)}$ for any k (the $\gamma(j)$'th entry of the eigenvector); let $\tilde{\pi}_{\gamma(j)}$ be equal to this value. Following this procedure for each signal release, after all information is released and disseminated, the final consensus will be

$$c^{(N)} = \sum_{j=1}^{N} \left(\prod_{k=j+1}^{N} \left(1 - \tilde{\pi}_{\gamma(k)} \lambda_{\gamma(k)}\right) \right) \tilde{\pi}_{\gamma(j)} \lambda_{\gamma(j)} s_{\gamma(j)} + \prod_{j=1}^{N} \left(1 - \tilde{\pi}_{\gamma(j)} \lambda_{\gamma(j)}\right) c^{(0)}$$

If a signal i is released in the very last information round, its influence on the final consensus will be $\tilde{\pi}_{\gamma(N)} \lambda_{\gamma(N)} = \tilde{\pi}_i \lambda_i$. If this same signal is released in any earlier information round $k < N$, it's weight $\tilde{\pi}_i \lambda_i$ will be multiplied with $\prod_{j=k+1}^{N} \left(1 - \pi_{\gamma(j)} \lambda_{\gamma(j)}\right)$. Since $\tilde{\pi}_{\gamma(j)} \in [0, 1]$ and $\lambda_{\gamma(j)} \in [0, 1]$, this additional argument suppresses the influence the signal has on the final consensus. Lowering k increases the number of arguments in the above product, further suppressing the weight of the signal.

Proposition 1. *Delaying the release of a signal increases its influence on the final beliefs.*

The proof follows from the above argument. This proposition, combined with a random assignment of participants, as is the case in the lab, leads to

Proposition 2. *Under a random assignment, earlier released signals have a lower expected* *influence on final beliefs.*

To prove the above proposition, it is sufficient to compare the expected weight on a signal released in round j and $j-1$. Let γ represent a particular assignment of agents to information release rounds, and $p(\gamma)$ the probability of such an assignment. In particular, $\gamma \in N_n$, where N_n represents the symmetric group (all the permutations of the N agents), and $p(\gamma) = \frac{1}{N!}$.[36] Agent j is characterized with $\tilde{\pi}_j$ and λ_j values, as well as signal s_j. Then, the expected difference between the weight of a signal released in round j and $j-1$ is

$$\mathbb{E}\left[\prod_{j=k+1}^{n}\left(1-\pi_{\gamma(k)}\lambda_{\gamma(k)}\right)\pi_{\gamma(j)}\lambda_{\gamma(j)}\right] - \mathbb{E}\left[\prod_{k=j}^{n}\left(1-\pi_{\gamma(k)}\lambda_{\gamma(k)}\right)\pi_{\gamma(j-1)}\lambda_{\gamma(j-1)}\right] =$$

$$\sum_{\gamma}\left(\prod_{j=k+1}^{n}\left(1-\pi_{\gamma(k)}\lambda_{\gamma(k)}\right)\left(\pi_{\gamma(j)}\lambda_{\gamma(j)} - (1-\pi_{\gamma(j)}\lambda_{\gamma(j)})\pi_{\gamma((j-1)}\lambda_{\gamma((j-1)}\right)\right)p(\gamma) \geq 0$$

The inequality follows from the random assignment with $p(\gamma) = \frac{1}{N!}$.

$\square$

On the other hand, note that as long as one signal is released in each information round, the influence of the common signal on the final consensus, $\prod_{i=1}^{N}(1-\tilde{\pi}_i\lambda_i)$, is unaffected by the timing of the signal release.

The arguments above rely on there being enough rounds between information release rounds for participants to stop updating their beliefs. Although we have limited rounds in between information release rounds, the data from graph figure 8, figure 9, and figure 10 show that beliefs stop evolving significantly two rounds(four rounds) after information release in the complete(ring) network. Thus, we believe there are sufficient rounds of communication for the analysis above to be a reasonable approximation of participants' behavior.

7.2.2 Sequential vs Simultaneous release and the Weight of the Common Signal

Simultaneous Information Arrival (Complete and Ring Network) Let M represent the network structure. In the absence of any weight adjustment, when information is released simultaneously, the weigh of the common signal on final beliefs will be $\left(1-\sum_{i=1}^{N}\pi_i\lambda_i\right)$. Where π_i corresponds to the i'th entry of the eigenvector, and λ_i is the weight agent i places on her signal.

For RSD agents, when information arrives simultaneously in the complete(ring) network for $z=1(z=3)$ rounds, the network matrix M is replaced with an alternative network matrix $\tilde{M}$. Let this modified network be $\tilde{M} = M + \tilde{\tilde{M}}$, and note that each row of $\tilde{\tilde{M}}$ sums

[36] $p(\gamma) = \frac{1}{N!}$ follows from the fact that each arrangement is equally likely. When assigning participants information arrival rounds, there is no selection based on any of their features.

to 0; that is, any weight increase $\tilde{\tilde{m}}_{i,j}$ must be offset by a weight decrease $\tilde{\tilde{m}}_{i,i}$. Define $\tilde{M}^{\infty} = M^{\infty}\tilde{M}$ in the complete network and $\tilde{M}^{\infty} = M^{\infty}\tilde{M}^3$ in the ring network. Note that each row in $\tilde{M}^{\infty}$ is identical, and the sum of each row is equal to one. The i'th entry within a row represents the i'th value in the eigenvector and, thus, the influence of agent i. Let $\tilde{\pi}$ represent the modified influence of agent i, then the weight of the common signal on the final beliefs will be

$$w_c^{sim} = 1 - \sum_{i=1}^{N} \tilde{\pi}_i \lambda_i$$

Again, since $\sum_{i=1}^{N} \tilde{\pi}_i = 1$, this is just a reshuffling of the initial influence vector.

Sequential Information Arrival (Complete Network) Now, consider RSD agents in a complete network when information arrives sequentially. After agent i receives her signal, this becomes apparent to all agents, and each agent places a higher weight on the action of agent i. Let $\hat{M}$ represent the additional weight on the action of agent i. That is, matrix $\hat{M}$ is equal to zero everywhere except for the entries in the i'th column. Let $\check{M}$ represent the decrease in the weights of all other agents. That is, $\check{M}$ is equal to zero on the i'th column and possibly positive on other columns. Note that the sum of each row of $\hat{M}$ is equal to the sum of each row of $\check{M}$, the additional weight placed on agent i is exactly equal to the weight subtracted from the other agents. Let $\tilde{g}$ be a vector equal to the previous consensus $c^{(i-1)}$ on $j \neq i$, and $\lambda_i s_i + (1 - \lambda_i)c^{(i-1)}$ on i. Then, after signal s_i is released and communication takes place, the new consensus will be

$$c^{(i)} = M^{\infty} \left(M + \hat{M} - \check{M} \right) \tilde{g} = M^{\infty}\tilde{g} + M^{\infty}\hat{M}\tilde{g} - M^{\infty}\check{M}\tilde{g}$$

This is a $N \times 1$ vector where each entry is equal to $c^{(i)}$, the new consensus formed after communication takes place. M^{∞} represents a matrix the rows of which are equal to the left eigenvector satisfying $\pi M = \pi$, corresponding to eigenvalue 1. That $M^{\infty}M = M^{\infty}$ follows from the alternative definition of the influence weights $\pi_i = \sum_j \pi_j m_{ji}$. The new consensus is then

$$c^{(i)} = \sum_j \pi_j \tilde{g}_j + \sum_j \pi_j \hat{m}_{ji} \tilde{g}_i - \sum_z \sum_k \pi_k \check{m}_{kz} \tilde{g}_z$$

Replacing all values $j \neq i$ $\tilde{g}_j$ with $c^{(i-1)}$ and $\tilde{g}_i$ with $\lambda_i s_i + (1-\lambda_i)c^{(i-1)}$

$$c^{(i)} = \left(\pi_i + \sum_j \pi_j \hat{m}_{ji}\right) \lambda_i s_i + \left(1 - \left(\pi_i + \sum_j \pi_j \hat{m}_{ji}\right) \lambda_i\right) c^{(i-1)}$$

Let $\tilde{\pi}_i = \left(\pi_i + \sum_j \pi_j \hat{m}_{ji}\right)$ and note that $\tilde{\pi}_i \in [\pi_i, 1]$, that is, since agents pay more attention to agent i when signal s_i is released, the weight it has on the new consensus is larger than the weight it would have if agents always placed fixed weights on their neighbors. Once more, if signals are released sequentially, one for each information release round, without loss of generality, assume that signal s_i is released in information round i, the final consensus will be

$$c^{(K)} = \sum_{i=1}^{N} \prod_{j=i+1}^{N} \left(1 - \tilde{\pi}_j \lambda_j\right) \tilde{\pi}_i \lambda_i s_i + \prod_{i=1}^{N} \left(1 - \tilde{\pi}_i \lambda_i\right) c^{(0)}$$

Thus, if all signals are released sequentially the weight on $c^{(0)}$ will be

$$w_c^{seq} = \prod_{i=1}^{N} \left(1 - \tilde{\pi}_i \lambda_i\right)$$

with $\tilde{\pi}_i \in [\pi_i, 1]$.

Sequential Information Arrival (Ring Network) In the ring network, when information is released sequentially, the network matrix is modified for three rounds. For exposition assume we are in information round one in which signal s_1 is received by agent 1. Denote the modification matrix $\tilde{M}$ as follows

$$\tilde{M}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \tilde{m}_4 & 0 & 0 & -\tilde{m}_4 \end{bmatrix} \tilde{M}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\tilde{m}_3 & \tilde{m}_3 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tilde{M}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & -\tilde{m}_2 & \tilde{m}_2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Agent four, who observed agent one, increases the weight on agent one by $\tilde{m}_4$. She can only do so by decreasing her own weight by the same amount. In the next round, agent three increases the weight on agent four, afterwards, agent two increases the weight on agent three, thus completing the information diffusion. Then, after three rounds of communication, the weight on s_1 will be

$$\begin{bmatrix} \lambda_1 m_{11}^3 \\ \lambda_1 \left(\tilde{m}_2 + m_{23}\right)\left(\tilde{m}_3 + m_{34}\right)\left(\tilde{m}_4 + m_{41}\right) \\ \lambda_1 \left(m_{11}m_{34}m_{41} + m_{33}\left(\tilde{m}_3 + m_{34}\right)\left(\tilde{m}_4 + m_{41}\right) + m_{34}\left(\tilde{m}_3 + m_{34}\right)\left(\tilde{m}_4 + m_{41}\right)m_{44}\right) \\ \lambda_1 \left(m_{44}^2\left(\tilde{m}_4 + m_{41}\right) + m_{41}\left(m_{11}m_{44} + m_{11}^2\right)\right) \end{bmatrix}$$

Note that for any value $\tilde{m}_i > 0$ the weight on s_1 is increased compared to SD agents, thus, for RSD agents, the weight on s_1 after communication takes place will be $\tilde{\pi}_1 \geq \pi_1$, and consequently the weight one the common signal will be lower. This holds for each signal s_i. Once more, the final consensus will be $c^{(K)} = \sum_{i=1}^{N} \prod_{j=i+1}^{N} \left(1 - \tilde{\pi}_j \lambda_j\right) \tilde{\pi}_i \lambda_i s_i + \prod_{i=1}^{N} \left(1 - \tilde{\pi}_i \lambda_i\right) c^{(0)}$, and thus the weight of the common signal on the final beliefs will be

$$w_c^{seq} = \prod_{i=1}^{N} \left(1 - \tilde{\pi}_i \lambda_i\right)$$

with $\tilde{\pi}_i \in [\pi_i, 1]$.

Simultaneous versus Sequential Let reaction refer to the weight shift towards the agent who just received information.

Proposition 3. *For a high enough reaction, the weight of the common signal on final beliefs is lower when information is released sequentially.*

To prove the above proposition recall from the analysis above that when information is released simultaneously $\sum_{i=1}^{N} \tilde{\pi} = 1$ continues to hold. Thus the weight on the common signal is bounded

$$\left(1 - \min_i \lambda_i\right) \leq w_c^{sim} \leq \left(1 - \max_i \lambda_i\right).$$

On the other hand, when information is released sequentially $\tilde{\pi}_i \in [\pi_i, 1] \; \forall i$, while the weight of the common signal on the final beliefs will be

$$w_c^{seq} \in \left[\prod_{i=1}^{N} \left(1 - \lambda_i\right), \prod_{i=1}^{N} \left(1 - \pi_i \lambda_i\right)\right].$$

As reaction increases ($\tilde{M}$ differs more and more from M), the $\tilde{\pi}_i$ weights increase towards 1. and w_c^{seq} monotonically moves towards it's lower limit, whereas w_c^{sim} moves within it's bounds, not necessarily monotonically. Since the lower bound of w_c^{seq} is lower than the lower bound of w_c^{sim}, a high enough reaction ensures that $w_c^{seq} < w_c^{sim}$.[37]

□

7.2.3 Comparisons

SEQ vs SEQ From section 7.2.1, we know that when releasing signals sequentially, the order of the particular signals does not affect the influence of the common signal on the final beliefs. However, the expected influence of a private signal decreases if the same signal is released in an earlier round. This implies that a sequential release of information, monotonically, from the lowest valued signal to the highest valued signal will lead to higher final beliefs than the reverse order.

SEQ vs SIM From section 7.2.2 we know that as long as agents pay sufficiently more attention to their neighbors who just received information, the weight on the common signal will be lower under sequential information release than under simultaneous information release. From section 7.2.1 we know that signals released in later rounds will have a higher expected influence on the final beliefs. Thus, shifting from simultaneous to sequential information release, the influence of the common signal shifts towards the individual signals, disproportionately more towards the latter released signals.

With these two statements in mind, when the common signal is lower than all the individual signals, it follows that the final consensus under sequential information release, in which signals are released from lowest to highest, will be higher than the final consensus under simultaneous information release.

On the other hand, when the common signal is higher than all the individual signals, it follows that the final consensus under sequential information release, in which signals are released from highest to lowest, will be lower than the final consensus under simultaneous information release.

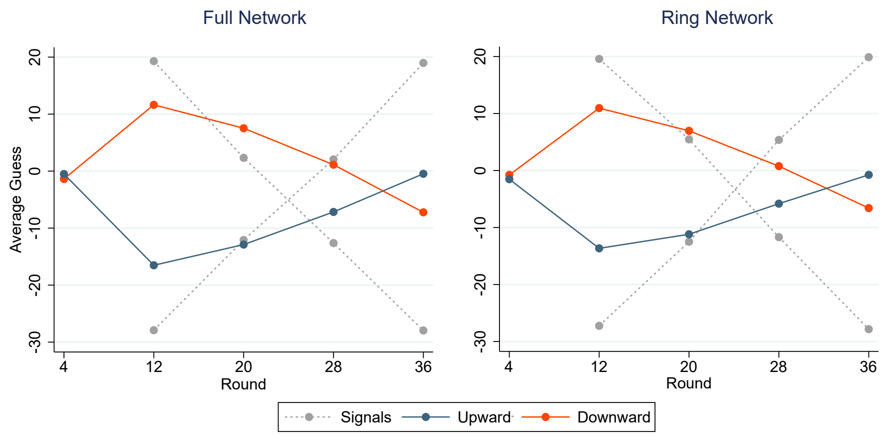
7.3 Data Analysis

7.3.1 Cross Plots by Treatment and Sequence Matches

Figure 6 presents a breakdown of the cross-plots analyzed in section 4.2 separately for each treatment.

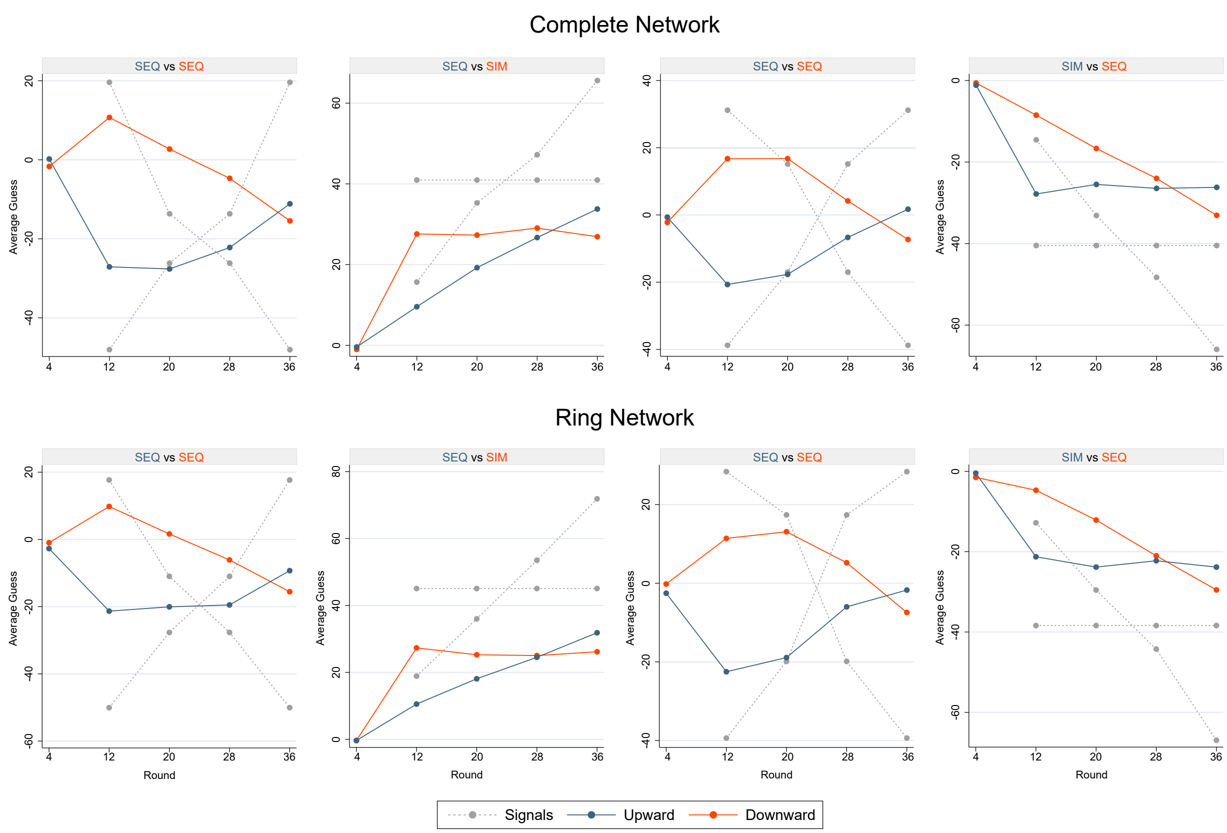
[37]Simulations reveals that the necessary reaction for the above to hold is rather low.

Figure 6: Observed Average Guesses By Treatment



[Figure 7](#) presents a breakdown of the cross-plots analyzed in [section 4.2](#) separately for each treatment and matched sequences.

Figure 7: Observed Average Guesses By Treatment and Sequence Matches

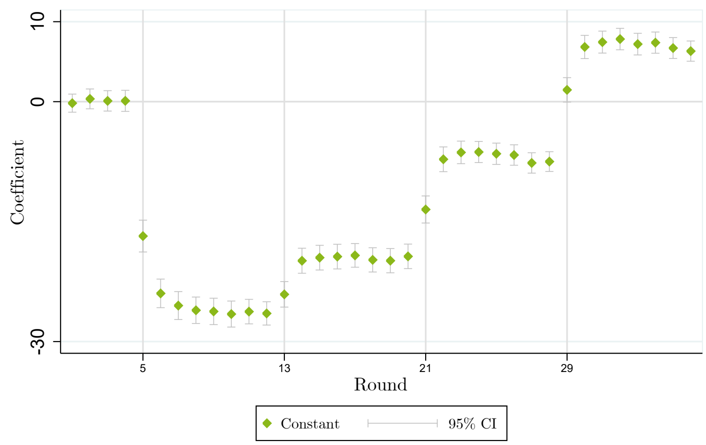


As can be seen, for each one of the matched sequences, for each one of the treatments, the upward sequences, the blue graphs, on average converge at higher average beliefs than the downward sequences, the orange graphs. Thus, we see this result not only on average across treatments but on average sequence by sequence and treatment by treatment as well.

7.3.2 The Evolution of Beliefs Across Rounds

Focusing on rounds in which information has arrived and disseminated, as we did on section 4.2, is informative with regard to the influence that information sequencing has on final beliefs. However, studying the evolution of beliefs across each round can help us grasp belief dynamics that we cannot notice if we focus only on rounds in which individual beliefs have roughly converged. Figure 8 plots the difference between upward and downward sequences across each round. Once more, to give each signal equal rounds to disseminate, we focus on rounds 1 through 36.

Figure 8: Difference - All Treatments and Sequences

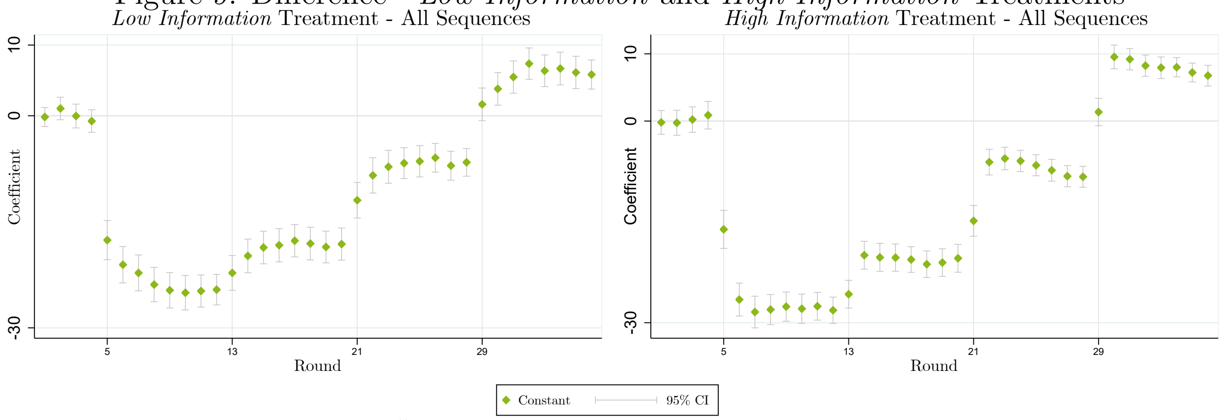


The graph above shows the average difference between the upward and downward sequences within each round. It also shows the 95% confidence interval of the estimated difference.

In the first four rounds, where the common signal is the only information the group has, the difference between the upward and downward sequences is not statistically distinguishable from 0. Between rounds 5 and 29, the difference between upward and downward sequences is negative. This is to be expected, as the downward sequences release higher signals earlier, whereas the upward sequences release lower signals earlier. That this difference is negative can be seen even from observing figure 3. The downward sequences (orange graphs) are typically above the upward sequences (blue graphs). However, when the very last signal arrives, which happens in round 29, both the upward and the downward sequences have released all the available information, which is in fact identical except for the timing of its arrival. What we see in figure 8 is that after all signals have been released (after round 29), the difference between upward and downward sequences becomes positive and statistically significant, as zero is not in the 95% confidence interval.

While we see that this difference is in the expected direction and statistically different from 0, to calculate this difference, we have pooled all sequences and treatments. It is of interest to see if this difference holds for the low-information treatment and the high-information treatments separately. Calculating the difference between upward and downward sequences and then regressing this difference on a constant, separately for the two treatments, roughly splits the data in half; thus, statistical power is lost. Nonetheless, as figure 9 shows, this difference continues to be statistically significant for each treatment.

Figure 9: Difference - *Low Information* and *High Information* Treatments

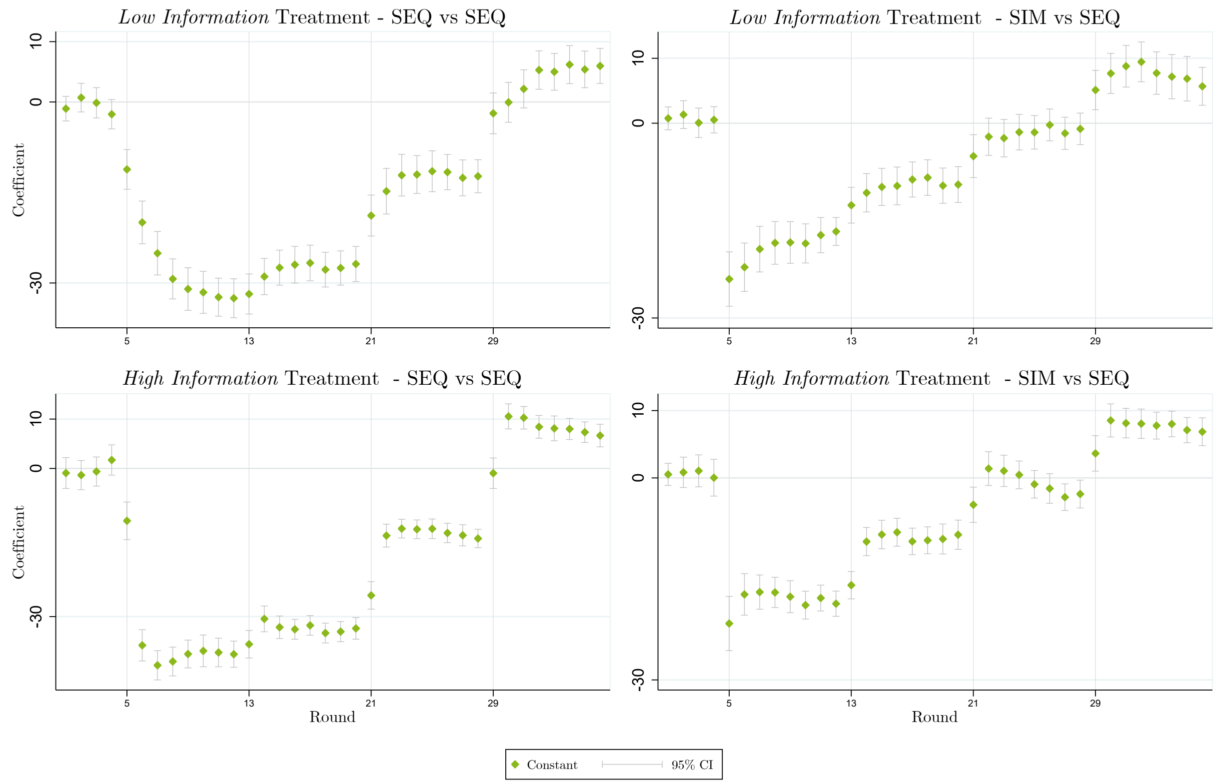


The graph above shows average difference between the upward and downward sequences within each round. It also shows the 95% confidence interval of the estimated difference.

Figure 9 further emphasizes the nature of information dissemination in the two treatments. In the low-information treatment, the network structure is a ring network. When a participant receives a signal, it takes one round for her neighbor to internalize her change, another round for her neighbor's neighbor to internalize her neighbor's change, and yet another round for the fourth participant to internalize the change. Thus, after a signal arrives, we can see that the average guess tends to change substantially in the next three rounds as the information disseminates across the network. We can see this in the rounds following information releases, such as rounds 5, 13, 21, and 29. In contrast, in the high-information treatment internalizing information is much quicker. Recall that the network structure here is a complete network. Thus, each time a participant receives a signal, the other three participants are affected in the very next round. Consequently, we see that the average guess changes substantially in information arrival rounds and one round after. In both treatments, after information disseminates, to a large extent, beliefs seem to stabilize.

Once more, while we see that the difference has the expected sign and is statistically significant for both treatments, it would be informative to see whether this is the case separately for the *SEQ* vs *SEQ* and the *SIM* vs *SEQ* matched sequences. Figure 10 breaks down the difference by treatment and sequence. By doing so, we slash the data roughly in half once again and, thus, further reduce the statistical power. Nonetheless, as can be seen in figure 10, the difference has the expected sign and remains significant in each case. Hence, not only is this effect prevalent on average when pooling the sequences and treatments, but it also holds on average within the sequence and treatment.

Figure 10: Difference - Each Treatment and each Sequence



The graph above shows the average difference between the upward and downward sequences within each round. It also shows the 95% confidence interval of the estimated difference.

As can be seen, after all information is released, and enough rounds of communication take place for information to disseminate (which is round 32 for the low-information treatment and 30 for the high-information treatment) the difference between upward and downward sequences is positive, and statistically significant.

References

Agranov, M., Lopez-Moctezuma, G., Strack, P., and Tamuz, O. (2021). Social learning in groups: an experimental study.

Agranov, M. and Reshidi, P. (2023). Disentangling suboptimal updating: Complexity, structure, and sequencing. *Working Paper*.

Ambuehl, S. and Li, S. (2018). Belief updating and the demand for information. *Games and Economic Behavior*, 109:21–39.

Angrisani, M., Guarino, A., Jehiel, P., and Kitagawa, T. (2018). Information redundancy neglect versus overconfidence: a social learning experiment. Technical report, cemmap working paper.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. *Handbook of Behavioral Economics: Applications and Foundations 1*, 2:69–186.

Brandts, J., Giritligil, A. E., and Weber, R. A. (2015). An experimental study of persuasion bias and social influence in networks. *European Economic Review*, 80:214–229.

Chandrasekhar, A. G., Larreguy, H., and Xandri, J. P. (2020). Testing models of social learning on networks: Evidence from two experiments. *Econometrica*, 88(1):1–32.

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

Choi, S., Gale, D., and Kariv, S. (2005). *Behavioral aspects of learning in social networks: an experimental study*, volume 13. Emerald Group Publishing Limited.

Corazzini, L., Pavesi, F., Petrovich, B., and Stanca, L. (2012). Influential listeners: An experiment on persuasion bias in social networks. *European Economic Review*, 56(6):1276–1288.

Dale, H. (1968). Weighing evidence: An attempt to assess the efficiency of the human operator. *Ergonomics*, 11(3):215–230.

Dasaratha, K. and He, K. (2019). An experiment on network density and sequential learning. *arXiv preprint arXiv:1909.02220*.

De Swart, J. and Tonkens, R. (1977). The influence of order of presentation and characteristics of the datagenerator on opinion revision. *Acta Psychologica*, 41(2-3):101–117.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121.

Edenborough, R. A. (1975). Order effects and display persistence in probabilistic opinion revision. *Bulletin of the Psychonomic Society*, 5(1):39–40.

Enke, B. and Zimmermann, F. (2017). Correlation neglect in belief formation. *The Review of Economic Studies*, 86(1):313–332.

Eyster, E., Rabin, M., and Weizsacker, G. (2015). An experiment on social mislearning. *Available at SSRN 2704746*.

Friedkin, N. E. and Johnsen, E. C. (1990). Social influence and opinions. *Journal of Mathematical Sociology*, 15(3-4):193–206.

Grether, D. M. (1992). Testing bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior & Organization*, 17(1):31–57.

Grimm, V. and Mengel, F. (2020). Experiments on belief formation in networks. *Journal of the European Economic Association*, 18(1):49–82.

Kallir, I. and Sonsino, D. (2009). The neglect of correlation in allocation decisions. *Southern Economic Journal*, pages 1045–1066.

Levy, G. and Razin, R. (2015). Correlation neglect, voting behavior, and information aggregation. *American Economic Review*, 105(4):1634–45.

Marks, D. and Clarkson, J. (1972). An explanation of conservatism in the bookbag-and-pokerchips situation. *Acta psychologica*, 36(2):145–160.

Mobius, M., Phan, T., and Szeidl, A. (2015). Treasure hunt: Social learning in the field. Technical report, National Bureau of Economic Research.

Mueller-Frank, M. and Neri, C. (2013). Social learning in networks: Theory and experiments. *Available at SSRN 2328281*.

Ortoleva, P. and Snowberg, E. (2015). Overconfidence in political behavior. *American Economic Review*, 105(2):504–35.

Peterson, C. R. and DuCharme, W. M. (1967). A primacy effect in subjective probability revision. *Journal of Experimental Psychology*, 73(1):61.

Peterson, C. R. and Miller, A. J. (1965). Sensitivity of subjective probability revision. *Journal of experimental psychology*, 70(1):117.

Pitz, G. F. and Reinhold, H. (1968). Payoff effects in sequential decision-making. *Journal of Experimental Psychology*, 77(2):249.

Reshidi, P. (2023). Information sequencing with naive social learning. *Working Paper*.

Roby, T. B. (1967). Belief states and sequential evidence. *Journal of Experimental Psychology*, 75(2):236.

Shanteau, J. C. (1970). An additive model for sequential decision making. *Journal of Experimental Psychology*, 85(2):181.

Shearer, E. and Mitchell, A. (2021). News use across social media platforms in 2020.