Behavioral and Structural Barriers to Information Aggregation in Networks*

Marina Agranov Ben Gillen Dotan Persitz

This version: September 3, 2025

"Imitation is not just the sincerest form of flattery - it's the sincerest form of learning." (George Bernard Shaw)

Abstract

We study how network architecture shapes learning dynamics in medium-sized groups using laboratory experiments. Participants are incentivized to guess the true state of nature based on two sources of information: (i) a private signal received before the game begins, and (ii) the past guesses of their immediate neighbors. We focus on identifying behavioral and structural frictions that impede successful learning. We find that subjects systematically under-react to new information—even in the Complete network—leading to persistent learning failures. This behavioral friction is exacerbated when the distribution of private signals is less informative. Additionally, in network positions where imitation is optimal, participants often fail to imitate better-informed neighbors (influencers). This under-imitation is too frequent to be explained as a rational response to influencers' own under-reaction to new information. Instead, evidence—including results from a novel intervention—implies that a lack of trust is the primary driver of this behavioral friction. As a result, networks with a single central fully connected node often perform poorly, even when most private signals are accurate. Beyond behavioral frictions, we define structural frictions as combinations of network topology and signal distribution that hinder aggregation even under fully rational behavior. Networks with centralized hubs or weakly connected cliques are especially prone to both behavioral and structural frictions in the presence of noisy signals. We argue that a dual framework of behavioral and structural frictions provides a sharper account of learning dynamics than standard Bayesian or naïve (DeGroot-style) models.

Keywords: Networks, Learning, Belief Formation, Information Aggregation, Under-Reaction, Under-Imitation, Laboratory Experiment.

JEL Codes: C91, C92, D03, D83, D85

*Agranov: California Institute of Technology, Pasadena, CA 91125 and NBER (e-mail: magranov@caltech.edu); Gillen: Robert Day School of Economics and Finance, Claremont McKenna College, Claremont, CA 91711 (e-mail: benjamin.gillen@claremontmckenna.edu); Persitz: Coller School of Management, Tel Aviv University, Tel Aviv (e-mail: persitzd@post.tau.ac.il). We thank Vered Shapira for outstanding design of the networks' images, Perfect Loop Technologies for providing support for the virtual experiments, and Zvika Messing, Ayelet Saly, May Melamed, Lior Sulkin, Michael Crystal and Hadar Shimon for excellent research assistance. We would like to thank Matt Elliott, Chaim Fershtman, Andrea Galeotti, Matthew Jackson, Pietro Ortoleva, Ariel Rubinstein, Leeat Yariv and participants in numerous seminars and conferences for helpful comments and suggestions. Persitz gratefully acknowledge financial support from the Israel Science Foundation (grant 1155/16), the Coller Foundation and the Henry Crown Institute of Business Research in Israel.

1 Introduction

In forming beliefs and making decisions, individuals integrate private information—drawn from their own knowledge and experiences—with insights gained by observing others' behavior. The extent to which an individual can observe others depends on their position in the social network, the overall structure of that network, and their beliefs about how their neighbors process the information they encounter.

To illustrate the process of belief formation through private and social information, consider the decision problem faced by an adolescent contemplating whether to wear a bicycle helmet. As governments increasingly promote cycling as an environmentally friendly mode of transportation, bicycle safety has become a crucial concern, making helmet use a key policy objective. While medical authorities advocate for mandatory helmet legislation and its enforcement,[1] environmentalists caution that such regulations may discourage cycling and lead to riskier riding behaviors (Molina-García et al. (2018)). An often-overlooked factor in adolescents' helmet use is the role of social learning. While a teenager may initially form beliefs about helmet safety based on information from parents, teachers, or online sources, those beliefs are subsequently updated through repeated interactions with peers. Each time the teenager joins a group bike ride, he observes whether others wear helmets, and over time, these observations accumulate. His decision on whether to wear a helmet in his next ride reflects both his prior beliefs and this evolving social evidence. This dynamic suggests that the observed correlation between peer helmet use and individual adoption may not stem solely from peer pressure—as is commonly understood—but could also reflect a process of information aggregation.[2] This pattern of belief updating—where private information is iteratively adjusted based on social observation—is common. Similar patterns arise in domains ranging from consumer purchases of products that are visible only to close acquaintances (e.g., home appliances, mattresses) to belief formation about workplace culture (e.g., wage renegotiation).

The theoretical literature on information aggregation in networks typically focuses on the asymptotic dynamics of societies of infinite size. This literature predominantly studies two benchmark approaches. One assumes fully Bayesian agents, who optimally extract all available information from their private signals, the network structure and the observed behavior of neighbors, often under the assumptions of myopia and the ability to communicate their beliefs (e.g. Gale and Kariv

[1] In 2023, the American College of Surgeons (ACS) updated its statement on bicycle safety and helmet promotion (American College of Surgeons (2024)), emphasizing that "Helmet use has been shown to significantly decrease the risk of both fatal and nonfatal head injuries. Based on these data, the ACS supports efforts to promote, enact, and sustain universal bicycle helmet legislation and enforcement." The ACS further reports that (i) more than 1,000 people die and 350,000 require emergency care annually due to bicycle-related injuries in the U.S.; (ii) in 2020, bicycle crashes resulted in 5.4 billion USD in medical costs and an additional 7.7 billion USD in lost productivity and lives lost; and (iii) helmets reduce the risk of head injury by 48%, traumatic brain injury by 53%, facial injury by 23%, and fatal injury by 34%. A 2022 technical report by the American Academy of Pediatrics provides similar statistics and also advocates for mandatory helmet laws (Lee et al. (2022)).

[2] The American Community Survey notes that "peer and adult companion helmet use is associated with increased bicycle helmet use by children." This pattern is typically attributed to peer pressure—for instance, Verlinde et al. (2024) argues that "peer pressure remains the primary barrier to helmet use among adolescents" (see also Lajunen and Räsänen (2004) and even Schelling (1973) on helmet adoption in hockey).

(2003), Acemoglu et al. (2011), Mueller-Frank (2013), and Mossel et al. (2015)). The other follows the DeGroot model (DeGroot (1974)) to posit naïve agents, who update their beliefs by simply averaging the beliefs of their neighbors (e.g. DeMarzo et al. (2003), Golub and Jackson (2010, 2012), and Acemoglu and Ozdaglar (2011)). Some studies adopt hybrid approaches, including frameworks where agents are neither fully Bayesian nor purely naïve (e.g., Bala and Goyal (1998), Goyal and Vega-Redondo (2005), and Mueller-Frank and Neri (2021)), or they allow for populations with heterogeneous updating rules (Chandrasekhar et al. (2020)). The key insight from this literature is that, under mild conditions on network structure, agents' inference abilities, and signal properties, beliefs in connected societies tend to converge to the truth.[3]

By contrast, experimental studies of information aggregation in networks have mostly focused on very small groups, typically involving 3 to 7 participants. A notable exception is Choi et al. (2023), who study three networks with 40 participants each.[4] Most of these experiments are urn-guessing games (Anderson and Holt (1997)) played over undirected networks, where each participant receives an initial noisy, informative signal about a predefined state and must repeatedly update their guess of the true state based on dynamically acquired information from their direct neighbors' actions (see Section 8.1). These experiments generally find deviations from Bayesian updating and require adjustments to match observed behavior to the naïve model.

We identify two key limitations in this literature. First, many real-world social networks are neither infinitely large nor as small as those studied in most experiments. Second, the prevailing focus on just two decision-making paradigms—Bayesian and naïve—limits the ability to account for deviations from rationality that are central to behavioral economics. By conducting laboratory experiments with relatively large networks, we aim to uncover both structural and behavioral frictions that influence individual-level information aggregation and the overall performance of different network structures. We show that expanding the analytical framework allows for a better understanding of both individual social learning behavior and network-level information aggregation outcomes.

We conduct a series of urn-guessing games played over undirected networks of size 18. Building on recent research in sociology and organizational science, we select six network architectures—alongside the Complete network—that incorporate at least one of two key features commonly observed in real-world networks: *hierarchy*, where a central node connects to all others, and *cohesiveness*, where subsets of nodes form tightly connected cliques. In each game, the true state is randomly determined

[3]Mobius and Rosenblat (2014), Golub and Sadler (2016) and Bikhchandani et al. (2024) provide excellent surveys of this literature. Bikhchandani et al. (2024) conclude that "An overarching conclusion ... is that egalitarianism in network structure, formalized in various ways, promotes information aggregation and welfare. This lesson holds across a variety of Bayesian, quasi-Bayesian, and heuristic settings."

[4]Several studies report experiments conducted in physical or virtual laboratories using networks of size between 15 and 50 nodes (see Berninghaus et al. (2002), Kirchkamp and Nagel (2007), Cassar (2007), Kearns (2012), Charness et al. (2014), Becker et al. (2017), Choi et al. (2017), Cardoso et al. (2020), Choi et al. (2024a,b), Bigoni et al. (2025)). However, none of these studies involve observational learning in settings where the network structure is explicitly revealed to subjects. As Choi et al. (2016) emphasize in the context of information aggregation: "Further experimental research is required to identify the type of bounded rationality ... [we have] ... to investigate how this updating varies with the size and complexity of the network, as the largest network explored so far has only 7 individuals."

by the computer to be either White or Blue, with probability 0.5 each. Each participant then receives a noisy private signal about the state that is correct with a probability of 70%. In the first round, participants make an initial guess about the state. From the second round onward, they observe their direct neighbors' guesses from all previous round and update their own guesses accordingly. The game concludes when no player changes their guess for three consecutive rounds. Participants are incentivized to guess correctly in each round.

We characterize the dynamics of each network and signal distribution under both the Bayesian and naïve models of information aggregation. These theoretical predictions serve two purposes: first, to demonstrate that neither model fully captures participants' behavior; and second, to establish benchmarks for identifying frictions in the information aggregation process. We say that a *structural friction* impedes information aggregation when it is common knowledge that all agents are myopically Bayesian and yet information aggregation may still fail. Our analysis identifies three network architectures in the experiment that are susceptible to such frictions. In one of them, a small subset of agents aggregates information on behalf of the entire network, without incorporating signals from outside the subgroup.[5] In the other two cases, the networks consist of two large cliques connected by only a few links. In cases where the signal distribution is such that the majority of signals in one clique differs from that in the other, the connectors—nodes with at least one direct neighbor from the other clique—must convey the relative strength of each majority through their actions alone. While correct aggregation is feasible in this setting, it poses a complex inference problem, akin to the classic "cheating spouses" logic puzzles studied in the literature on common knowledge. We say that in these cases we encounter *surmountable structural friction*, that is, a situation in which successful information aggregation is theoretically possible but requires an exceptionally high level of sophistication from decision-makers.

We evaluate network performance using an Aggregate Learning Index (ALI), which measures the extent to which participants revise incorrect private signals into correct final guesses. While the Complete network exhibits the highest level of information aggregation, its performance still falls short of the predictions generated by both the Bayesian and naïve models. *Single-Aggregator* networks—where one node is connected to all others—perform comparably to the Complete network when the signal distribution is close to uniform, but, surprisingly, fail to improve as the number of correct signals increases. As a result, these networks often fail to aggregate information effectively, even when a large majority of participants receive correct private signals. In contrast, *Cluster(s)* networks—where the network includes one or two large cliques—are responsive to the overall signal quality. They match the performance of the Complete network when most signals are correct but frequently fail when the signal distribution is close to a tie.

To unpack the sources of these aggregation failures, we analyze behavior at the positional level. We find that most of the informational dynamics unfold in the first three rounds. Any reasonable model of myopic agents predicts that participants rely on their private signals in the first round.

[5] This result mirrors the "royal family" argument proposed by Bala and Goyal (1998) for directed networks. It may also be understood as an instance of the related "Majority Illusion" (see Lerman et al. (2016) and Jackson (2019)). Bikhchandani et al. (2024) argue that such networks lack egalitarianism (see Footnote 3).

Indeed, in more than 92% of cases, participants' initial guesses align with their private signals. Before making their second-round guess, participants observe the guesses of their direct neighbors. Both the Bayesian and the naïve models predict that participants will follow the majority of signals inferred from their neighbors' guesses, combined with their own private signal. However, we find that across all networks and positions, participants who belong to the local minority systematically *under-react to new information* from their neighbors' first-round guesses and tend to stick with their own initial guess. Notably, this friction diminishes as the strength of the evidence in favor of the majority signal increases. This behavior is consistent with well-documented under-reaction to new information, observed across a wide range of settings (Benjamin (2019)).

In round three, indirect information from non-neighboring nodes becomes accessible. According to Bayesian benchmarks, agent i should imitate their direct neighbor, agent j (the influencer), if agent j is strictly better informed than i and all of i's other neighbors. However, we observe systematic *under-imitation* across all networks and positions when the second-round guess of the influencer differs from that of the potential imitator. This finding is consistent with prior experimental studies in sequential social learning settings, where participants tend to imitate predecessors less frequently than optimal when imitation requires going against their private signal (Weizsäcker (2010), Ziegelmeyer et al. (2013)). We show that imitation rates in our setting are far too low to be explained solely as a rational response to under-reaction to new information. In addition, we establish that under-imitation is not a result of naïve behavior. Our analysis reveals that these two behavioral frictions are closely related. In particular, we identify behavioral patterns suggesting that under-imitation may reflect a lack of confidence in the informational value of the influencer's behavior. First, imitation rates increase when the local majority supports the influencer rather than the potential imitator. Second, imitation becomes more likely when the influencer is observed to switch their guess between the first and second rounds—indicating that they are not under-reacting to new information.

To further examine these two behavioral frictions—and their relationship—we conduct a targeted intervention designed to mitigate under-reaction to new information. We focus on the One Gatekeeper network, in which 9 participants form a clique and the remaining 9 "leafs" are each connected solely to the same clique member, the "aggregator." In each game of the new treatment, non-aggregators receive the same private signal as in the original sessions, whereas the aggregator receives no signal—a design feature known to all participants. This intervention significantly improves performance: ALI increases; over 70% of aggregation errors in the original sessions—among aggregators initially in the minority—are eliminated; and imitation rises among participants who disagree with the aggregator—particularly leafs and second-round minority clique nodes. Together with prior evidence, these results suggest that the observed low levels of imitation stem from a lack of trust in the aggregator's second-round behavior, which reflects their under-reaction to new information. When a clear reason for trust is introduced (i.e., the aggregator has no private signal), imitation increases accordingly.

Our final step is to link the micro-level behavioral frictions and the structural frictions to aggregate-level performance. The complete network under-performs due to under-reaction to new information—an effect that intensifies as signal quality declines. *Single Aggregator* networks suffer from compounded behavioral frictions: under-reaction by the aggregator and under-imitation by the other participants. Moreover, the aggregator's under-reaction appears insensitive to the size of the first-round minority—possibly due to the absence of monitoring—so performance does not improve even as signal quality increases. Finally, performance in *Cluster(s)* networks is hindered by both structural and behavioral frictions. These frictions are particularly damaging when signal quality is weak, resulting in poor performance. However, they are largely neutralized when signal quality is high, allowing these networks to match the performance of the complete network under strong signals.

The remainder of the paper is organized as follows. Section 2 describes the experimental design. Section 3 presents the Bayesian and naïve benchmarks. For each benchmark, we outline general insights into predicted behavior and apply these predictions to the network structures used in our experiment. Section 4 evaluates network-level performance using the ALI and shows that failures in information aggregation cannot be explained by structural frictions alone. This section also examines the performance of the naïve model and the dynamics of the information aggregation process. Section 5 analyzes positional behavior in the first three rounds and in the final round of the game. Section 6 introduces our intervention and reports its outcomes. Section 7 then connects these macro-level results to the behavioral and structural frictions identified earlier and highlights the network positions that consistently perform well or poorly. Section 8 concludes with a discussion of: (i) connections to the experimental literature; (ii) the (in)-compatibility of alternative models, beyond the Bayesian and naïve benchmarks; and (iii) the decision-making procedures used by participants. Formal statements, proofs, and robustness exercises are provided in the theoretical and empirical appendices.

2 The Experiment

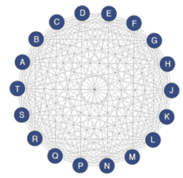
Here we describe the networks we study and present the details of the experimental protocol.

2.1 Networks

Recent research in sociology and organizational science has identified hierarchy and cohesiveness as two distinct, fundamental features of real-life networks (e.g. McFarland et al. (2014) and Bernstein et al. (2023)). Our experimental design focuses on undirected networks exhibiting at least one of these features. We operationalize hierarchy by introducing a central node connected to all others, and cohesiveness through the inclusion of cliques (i.e., fully connected subsets of nodes). Furthermore, most networks in our design are pairwise stable under simple variations of the well-known strategic network formation game described in the connections model introduced by Jackson and Wolinsky (1996).[6] Following this rationale, we chose the following seven networks, each with 18 members.

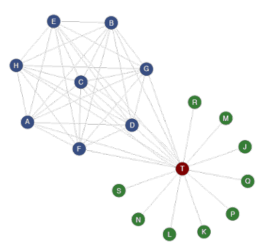
[6]For an overview, see Chapter 6 in Jackson (2008). For the specific network structures, refer to Jackson and Wolinsky (1996), Jackson and Rogers (2005), and Persitz (2010).

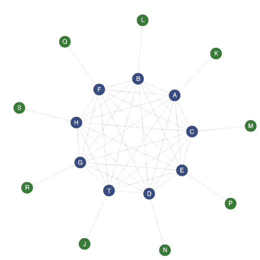
- *The Complete network* is a fully connected set of 18 nodes (a clique). This network represents the upper bound for information aggregation when there are no connectivity restrictions.



Complete

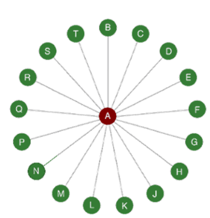
- *Core periphery networks* are characterized by two distinct types of positions: the core and the periphery. In our design, the core consists of 9 members who are directly connected to each other, forming a clique. The periphery comprises 9 members, each connected to a single core member. We examine two variants of this structure. In the Symmetric Core Periphery network, each peripheral node is linked to a distinct core member, resulting in a balanced distribution of connections between the core and periphery (Bala and Goyal (1998) refer to a similar directed network as a "Royal Family"). In the One Gatekeeper network all peripheral nodes are connected to a single core member, referred to as the "Gatekeeper". Note that the Gatekeeper is connected to all others nodes. This creates a hierarchical structure where the Gatekeeper serves as the sole intermediary between the core and periphery.

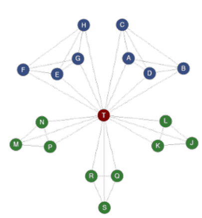




One Gatekeeper Symmetric Core Periphery

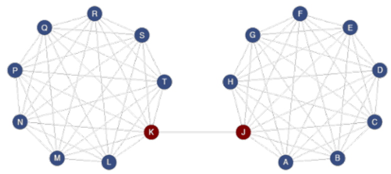
- *Hub and Spokes networks* feature two types of positions: the hub, which is connected to all other agents (as the Gatekeeper in the One Gatekeeper network), and the spokes, which may or may not be connected with each other. The two chosen networks feature disconnected spokes (Star) and locally segregated neighborhoods of spokes (Connected Spokes).



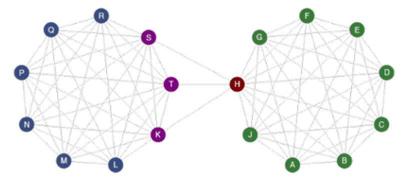


Star Connected Spokes

- *Multiple cliques networks* include two cliques of size 9 that are sparsely connected. The two chosen networks differ in the number of connections between the cores. The Two Cores with One Link network exhibits a single connection between two *connectors*, one from each clique. The Two Cores with Three Links network features a single connector in one clique that is directly connected to three nodes in the other clique.



Two Cores with One Link



Two Cores with Three Links

2.2 Experimental Protocol

Main game. In each session, a group of 18 participants plays 10 repetitions of the main game with one of the network structures described above. At the beginning of the main game, participants are randomly assigned a position in a network and observe its visual representation depicting all the connections between the players. At the same time, nature randomly determines the state, which is either WHITE or BLUE with equal chance. The state is fixed for the duration of the game and is shared by all eighteen players. Before the first round, each player gets a partially informative private signal about the state; the signals are conditionally independent and are correct with a probability 70%. After observing private signals, players are prompted to guess the state. In round two, and all the subsequent rounds, players can observe guesses made by their direct neighbors in all previous rounds and guess the state again. The information about neighbors' guesses is summarized in an intuitive way on the screen and is accessible at any point in time throughout the game (Figure 1). That is, we implement perfect recall by providing participants with all the information they have observed at any point in the game, and allowing them to quickly access this information. We chose this intuitive, comprehensive, accessible, and visual interface to isolate the effect of network architecture on learning, minimizing potential confounds such as imperfect memory or incomplete information.

Whereas most information aggregation experiments (e.g. Choi et al. (2005, 2012) and Mueller-Frank and Neri (2015)) impose a fixed predefined number of rounds, we chose not to do so in order to avoid "last-round effects" and to ensure that information aggregation is exhausted. In our design, the game ends in one of two ways. First, the game ends when all eighteen players submit the same guesses in three consecutive rounds. These do not have to be the same guesses across players, but it has to be the case that no player changes her mind in the last three consecutive rounds.[7] Second, if

[7]The diameter of a network is the longest shortest path between any two agents. The diameter is considered to be a natural baseline for the number of periods required for information to flow through the entire network. The largest diameter in our networks is three, which dictates our choice of three rounds of "inactivity" as an indication that a subject has exhausted her learning potential.

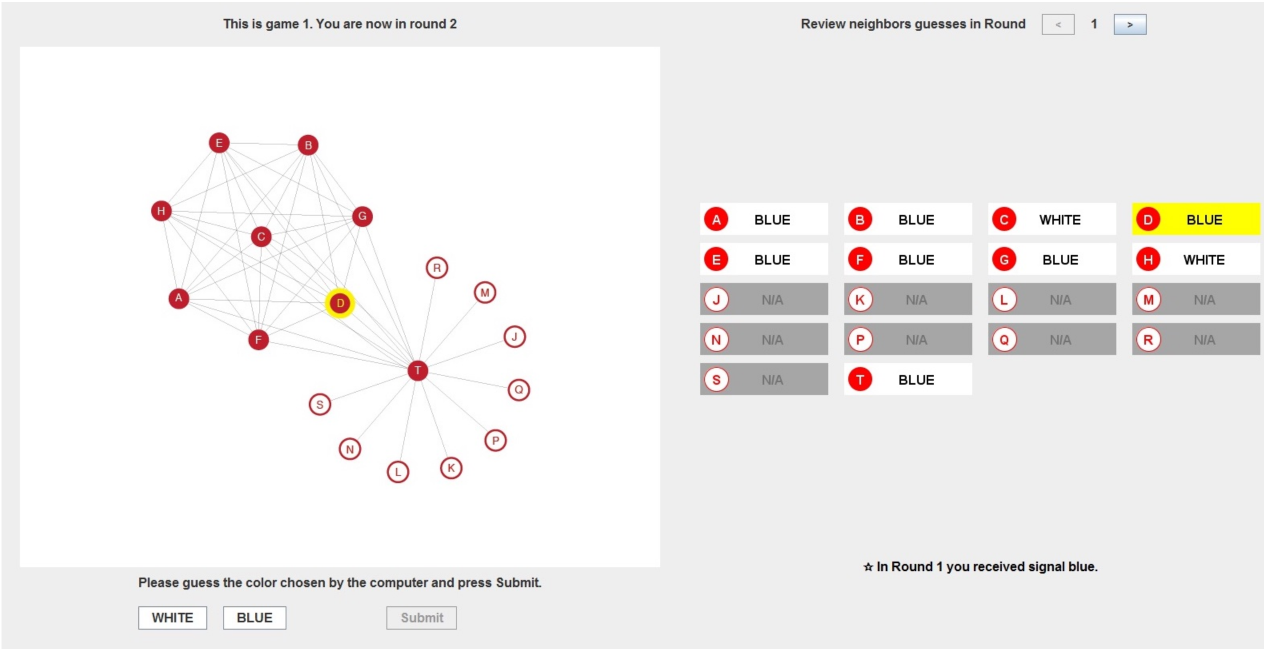


Figure 1: Screenshot of Beginning of Round 2, One Gatekeeper Session

Notes: The screenshot, taken at the beginning of round 2 of game 1, displays the experimental interface used in the One Gatekeeper network session as seen by Player D. On the left, the network configuration (fixed throughout the game) is shown: the subject's role (Player D) is highlighted with a yellow circle, while Player D's direct neighbors are marked with red-filled circles; all other participants appear as hollow circles. Below the network image, participants interact with decision buttons. When either the WHITE or BLUE button is selected, the SUBMIT button becomes active, allowing the participant to submit their guess. On the right, a table summarizes the participant's own guesses and those of their direct neighbors from all previous rounds—the current view shows the guesses from round 1. Navigation arrows above the table enable participants to review their guessing history across rounds. Additionally, the private signal, which remains unchanged throughout the game, is displayed beneath the history table.

the game reaches round 50, then there is a 50% chance that each next round is the last one. We use this ending procedure as a safeguard against situations in which some players continue to switch endlessly.[8] Participants are rewarded for the accuracy of their guesses. Specifically, at the end of the experiment, one game is randomly selected for payment. Then, one round of this chosen game is randomly selected for payment (Azrieli et al., 2018). A participant receives $20 if she guesses the state correctly in this chosen round and $5 if her guess is wrong.[9] We provide the instructions for the One Gatekeeper network game in Section A.1 of the Empirical Appendix.

Other experimental details. At the end of the experiment, subjects complete several incentivized short control tasks. These include the elicitation of risk attitudes (Gneezy and Potters (1997)) and their tendency to probability match—i.e., to choose an action with a frequency equal to

[8]This game termination scheme was activated in only 31 of 410 games. Most of these (22) occurred in the first two games of a session. Chandrasekhar et al. (2020) use random termination as the only game ending scheme.

[9]The payments in the experiment that took place in Israel were 60 NIS for a correct guess and 15 NIS for an incorrect guess.

	# of sessions and their location				# sessions	# subjects
	UCI	UCSD	TAU	OSU		
Complete	2	2	1	0	5 sessions	106 subjects
Star	2	1	1	2*	6 sessions	121 subjects
One Gatekeeper	2	2	2	0	6 sessions	122 subjects
Symmetric Core Periphery	6*	0	0	0	6 sessions	120 subjects
Connected Spokes	2	2	2	0	6 sessions	122 subjects
Two Cores with One Link	2	2	2	0	6 sessions	120 subjects
Two Cores with Three Links	2	2	0	2*	6 sessions	122 subjects

Table 1: Experimental Sessions

Notes: The number of sessions conducted at each location for each network is reported. In the last two columns we summarize the total number of sessions per network and the total number of participants per network. * indicates sessions that were conducted online due to the closure of physical labs during COVID-19 times.

the probability of that action being optimal, a clearly suboptimal heuristic.[10] Section A.3 of the Empirical Appendix provides these tasks as well as an additional short survey. Measures derived from these tasks are used as controls in the individual-level analysis (see Section A.4 of the Empirical Appendix for details).

We conducted 47 sessions. Each session lasted on average 90 minutes and the average total payment was $23.9, including $7 participation fee. To make sure that the rules of the game were common knowledge, the experimenter read the instructions out loud and all participants had to complete a comprehension quiz and answer all the questions correctly (see Section A.2 of the Empirical Appendix). Because of the large number of subjects required for our experiments, we conducted the experiment at four different locations: the University of California in Irvine, the University of California in San Diego, Ohio State University, and Tel Aviv University.[11] The early sessions were conducted using the Multistage software developed at Social Science Experimental Laboratory in Caltech. Due to incompatibility issues between Multistage and newer versions of JAVA we switched to oTree while keeping the interface and the procedures identical (Chen et al., 2016). Finally, due to the COVID-19 pandemic, the last 10 sessions were conducted online rather than in a physical lab. The subject pool for the online sessions was the same as in the physical lab (undergraduate students in one of the four universities) and we kept the protocol identical between the two types of sessions (for a comparison see Section A.6 of the Empirical Appendix and Rigotti

[10]Probability matching has been documented across various domains (see Humphreys (1939), Grant et al. (1951), Siegel and Goldstein (1959), Loomes (1998), and Rubinstein (2002)). See Footnote 7 in Choi et al. (2012) for an example of probability matching in an information aggregation experiment. Rivas (2013) provides a recent account of the connection between probability matching and reinforcement learning. Agranov et al. (2023) study experimental methods for mitigating probability matching in the laboratory.

[11]We conducted a few pilot sessions at the Experimental Economics Laboratory in Ben Gurion University of the Negev. The goal of these sessions was mainly to test the functionality of the software. These pilot sessions had different network structures in each game of a session, which resulted in much more noisy behavior than when participants play the same game 10 times.

et al. (2023)). Table 1 summarizes the experimental sessions.[12]

3 Theoretical Benchmarks

In Section 3.1, we introduce the theoretical setting. Sections 3.2 and 3.3 then present two benchmark models. The first assumes that agents are myopic Bayesian utility maximizers and that this is a common knowledge. The second follows the naïve-learning framework of DeGroot (1974). In Section 3.4, we apply these predictions to the network structures used in our experiment. Throughout, for clarity and conciseness, we state the main results informally with their intuitive explanations. Formal statements and proofs are relegated to Sections A and B of the Theoretical Appendix.

3.1 Belief Formation over Communication Networks: Theoretical Setting

Consider an undirected network $G =< N, E >$ where $N = \{1, 2, \ldots, n\}$ is the set of agents and E is the set of edges. The edge $ij \in E$ indicates that agents i and j are directly connected. $B(i) = \{j : ij \in E\}$ denotes the set of agent i's direct neighbors with cardinality $b(i)$. We assume that there are no isolated agents, i.e., $\forall i : b(i) > 0$. A subset of agents, $C \subseteq N$, forms a clique in G if (i) each pair of agents in this subset is directly connected, $\forall i, j \in C : ij \in E$, and (ii) there is no other agent that is connected to all clique members, $\forall k \in N \backslash C, \exists i \in C : ik \notin E$.

There are two equally probable states of nature, $\omega \in \{\text{WHITE}, \text{BLUE}\}$. Every agent i receives a signal $s(i) \in \{w, b\}$. Conditional on the realized state ω, signals are independently and identically distributed across agents and match the true state with probability $q \in \left(\frac{1}{2}, 1\right)$. If $\omega = \text{WHITE}$ then $s(i) = w$ with probability q and $s(i) = b$ with probability $1 - q$; similarly, if $\omega = \text{BLUE}$ then $s(i) = b$ with probability q and $s(i) = w$ with probability $1 - q$. The signals' accuracy parameter q is common knowledge, but the state of nature ω is unknown to the agents.

The belief formation dynamics begins after the state is realized and the agents receive their private signals. In each round $t \in \{1, 2, \ldots\}$, agent i chooses an action $a_i^t \in A = \{\text{W}, \text{B}\}$. In round 1, the only information available to the agent is her own private signal. In later rounds, before making a choice, each agent observes the past actions of her direct neighbors. We assume perfect recall: in every round, they can observe their own private signal, their past actions, and the complete actions' history of their direct neighbors. In each round, the agent's objective is to guess the state that corresponds to the majority of private signals.[13] An agent's payoff is 1 for a correct guess and 0 otherwise. In case of a tie, any guess is considered to be correct.

[12]In each session, more than 18 participants were recruited. At the start of each game, 18 subjects were randomly assigned to play, while the rest served as *observers*. No subject was assigned the observer role in two consecutive rounds. Observers chose one network position whose payoff they would receive if the game was selected for payment. Their information was only the network structure. Thus, they were incentivized to select the position they perceive as the most desirable. We analyze observer choices in Agranov et al. (2025).

[13]In the experiment, the goal of each subject was to guess the correct state of nature. Note that in a finite setting, it might happen that the realized majority of signals differs from the true state. This event is rare given our experimental parameters, $n = 18$ and $q = 0.7$; theoretically, it happens about 2% of the time (see Table 2 in Section B of the Empirical Appendix) and empirically it happened in 10 out of 410 games. To reconcile the theoretical setting with our experimental setup, for the data analysis, in those 10 cases, we redefine the state to be the majority of signals.

3.2 The Myopic Bayesian Model of Belief Formation over Communication Networks

Assume that all agents are myopic Bayesian utility maximizers, and that this is common knowledge. That is, each agent knows her own signal, takes a myopically optimal action in each round, and dynamically forms beliefs about the other $n-1$ signals (see discussion in Golub and Sadler (2016)). In this section, we focus on general insights provided by the myopic Bayesian model of belief formation over networks. Specific predictions for the networks used in our experiment are presented in Section 3.4.

Before proceeding, we briefly discuss the implications of the model's assumptions. First, assuming myopia is restrictive.[14] In some network structures and in some signals' distributions, it may be dynamically optimal to take actions that are not optimal from a myopic perspective.[15] Second, due to the myopia assumption, the model describes a decision problem rather than a strategic game. This is because there are no payoff externalities: each agent's preferences depend only on her own actions and the distribution of initial signals. Without myopia, dynamic considerations could turn the setting into a game of strategic information revelation. Third, the model implies that each agent believes others act according to myopic Bayesian reasoning. This assumption may be violated in experimental settings, where subjects might identify behavior inconsistent with Bayesian inference.[16] Finally, agents hold no prior over the behavior of others who are indifferent between actions based on their histories (see Footnote 42). As a result, in some networks and for some signals' distributions, the model does not yield a unique prediction.

In the first round, since the signals are informative, all agents choose the action that corresponds to their private signal (Lemma 1 in Section A of the Theoretical Appendix). Due to the common knowledge that everyone is myopic Bayesian, the agents know that their direct neighbores' disclose their private signals through their first-round actions (Lemma 2 in Section A of the Theoretical Appendix) which are observed before taking the second round's action. This feature of the model has important implications for subsequent play: any additional information agents acquire after the first round pertains only to their beliefs regarding the signals of non-neighbors ($N \backslash (\{i\} \cup B(i))$).

In the second round, the optimal behavior entails choosing the action that aligns with the majority of first-round actions among an agent's direct neighbors—which, given first-round behavior, corresponds to the majority of private signals in the agent's local neighborhood (Lemma 3 in Section A of the Theoretical Appendix).[17]

Optimal behavior from the third round onward is more difficult to characterize in general, as it depends on the network structure, the agent's position within it, and the realized distribution of

[14] Myopia is typically justified either by assuming each node represents a continuum of agents (e.g., Gale and Kariv (2003)), or by imposing strong discounting.

[15] For example, agent i might wish to study how her actions influence those of her neighbors, using these reactions to better infer their information and beliefs.

[16] For example, selecting an action that contradicts the majority in the Complete network, or switching actions in the absence of new information. See Chandrasekhar et al. (2020) for a related discussion of mixed models in which some agents are myopic Bayesian and others are naïve.

[17] Lemma 4 in Section A of the Theoretical Appendix shows that, due to the common knowledge that all agents are myopic Bayesian utility maximizers, an agent can deduce some information on their neighbors' neighbors after observing the second-round guesses.

signals. However, Agranov et al. (2024) provide a characterization of network positions for which it is optimal to imitate a selected neighbor. We say that agent i imitates agent j when $\forall t > 2 : a_i^t = a_j^{t-1}$.[18] Intuitively, Agranov et al. (2024) show that imitation is optimal for agent i when one of her neighbors, j, possesses strictly superior information relative to i and to each of her other neighbors. Formally, agent j is said to be better informed than agent i if $B(i) \cup \{i\} \subset B(j) \cup \{j\}$, denoted by $j \rhd i$.[19] Under the myopic Bayesian model, such an informational advantage implies that j has nothing to learn from i's inferences. That is, if $j \rhd i$, then j should ignore i's actions after observing her private signal in the first round (Lemma 5 in Section A of the Theoretical Appendix). Moreover, if j is also better informed than all of i's other neighbors, then j possesses every piece of information that might reach i through alternative paths, before i receives it herself. Hence, by imitating j, agent i does not forgo any potential future information. Define the set of i's neighbors who are strictly better informed than i and all her other neighbors as $C(i) = \{j \in B(i) | \forall k \in \{B(i) \backslash \{j\}\} \cup \{i\} : j \rhd k\}$. Proposition 1 from Agranov et al. (2024), re-stated below, shows that if $C(i)$ is non-empty, it must be a singleton containing the unique neighbor whom i should imitate—referred to as the "influencer".

Proposition 1. *Let $i \in N$. Then, $C(i)$ is either empty and imitation could lead to sub optimal behavior by agent i or it is a singleton, $C(i) = \{j\}$, and $\forall t > 2 : a_i^t = a_j^{t-1}$ is optimal for agent i.*

In Section 3.4, we show that many positions in our experimental networks satisfy this characterization. Therefore, under the myopic Bayesian model, imitation is the optimal strategy for agents in those positions.

3.3 The Naïve Model of Belief Formation over Communication Networks

The influential model of naïve belief formation introduced by DeGroot (1974) assumes that in each period $t > 1$, agents update their beliefs by taking a weighted average of their own belief and the beliefs of their direct neighbors from period $t - 1$.[20]

We focus here on the simplest version of the model, in which each agent assigns equal and fixed weights of $\frac{1}{b(i)+1}$ to her own belief and to each of her $b(i)$ neighbors. In our binary setting—where the state, signals, and actions are binary—this rule is sometimes referred to as the DeGroot action model. It implies that agents guess according to their private signal in the first round, and follow the local majority of period $t - 1$ in each subsequent round $t > 1$. In the case of a tie, either action is permissible. For a formal statement see Definition 1 in Section A of the Theoretical Appendix. In Section 8.2, we revisit our experimental findings in light of more complex forms of naïve updating discussed in the literature.

[18] The imitation principle in Bala and Goyal (1998) uses imitation in settings that include feedback, as a possible strategy, and infers that the agents' payoffs should be similar in equilibrium. Notably, that principle does not account for the optimality of imitation.

[19] This definition implies that: (a) i and j are direct neighbors; (b) there exists at least one agent $k \in B(j)$ such that $k \notin B(i)$; and (c) j has a finer information structure than i. According to Green and Stokey (1978), this means that j is more informed than i, and by Blackwell (1953), she therefore has a higher expected payoff.

[20] For surveys, see Section 8.3 in Jackson (2008) and Section 3 in Golub and Sadler (2016). The Naïve belief formation model can be represented as a Quasi-Bayesian model of Mueller-Frank and Neri (2021) with a specific functional form.

The behavior of naïve agents in the first two rounds is straightforward: they report their private signal in the first round and follow the majority of their local neighborhood's first-round guesses in the second. Combined with the discussion in Section 3.2, this implies that naïve agents are behaviorally indistinguishable from myopic Bayesian agents during the first two rounds, since both rely on their private signal in round 1 and on the local majority in round 2.

A key feature of collective naïve behavior is the rapid stabilization of beliefs within highly cohesive groups (see Morris (2000)). Informally, consider a clique C in network G, and let $\hat{C}$ denote the subset of members in C for whom the observed first-round majority in their local neighborhood coincides with the first-round majority in C. Lemma 6 in Section A of the Theoretical Appendix shows that if $\hat{C}$ is sufficiently large, then its members follow the clique's majority from the second round onward and maintain it indefinitely with no regard to information outside the clique.[21]

3.4 Predicted Dynamics Network-by-Network

In this section, we use both the Bayesian model and the naïve model to derive predictions regarding the dynamics of guesses across the seven networks implemented in our experiment. For clarity and accessibility, we present informal statements in the main text; formal versions and corresponding proofs are provided in Section B of the Theoretical Appendix. The section concludes with Table 2, which offers a concise summary of the theoretical predictions.

The Complete network (Result 1 in Section B of the Theoretical Appendix). In the Complete network, there are no connectivity restrictions as all agents are directly connected. Under both the Bayesian and the naïve models, if there is no tie in the signals' distribution, agents will converge to the correct guess already in the second round, with no further switches. In the case of a tie, neither model yields a prediction.[22]

Networks with a single aggregator (Results 2, 3, 4 and 5 in Section B of the Theoretical Appendix). Three networks— the One Gatekeeper, the Star, and the Connected Spokes—feature a single aggregator, that is, a single node connected to all other nodes.[23] Under the myopic Bayesian model, and assuming no tie in the signal distribution, all agents converge to the correct guess by the third round, with no further switches thereafter. In round 1, each agent reports their private signal, allowing the aggregator to observe the full signal distribution. In round 2, the aggregator

[21] Lemma 6 analyzes the behavior of naïve agents in cliques, whereas Proposition 1 in Chandrasekhar et al. (2020) examines naïve consensus behavior in clans, and Proposition 2 in Mueller-Frank and Neri (2021) highlights failures of information aggregation by naïve agents. Despite these differences, all three results rest on a common principle in the DeGroot action model: highly connected naïve agents are unlikely to change their actions beyond the first few rounds. A similar idea appears in Golub and Jackson (2012) in the context of the DeGroot belief model, where Proposition 3 shows that the rate of convergence of beliefs in a multi-type random network is determined by the rate of learning across cohesive groups, rather than within them.

[22] The naïve model, however, predicts that if a majority is formed in period t, all agents follow the majority guess in all subsequent periods.

[23] Formally, a network with a single aggregator contains a unique node i, referred to as the aggregator, such that $B(i) = N\backslash\{i\}$, while for any other node, $j \neq i$, there exists at least one node $k \notin \{i, j\}$ such that $jk \notin E$, that is, $B(j) \subset N\backslash\{j\}$.

reports the majority signal (if one exists), while non-aggregators follow their local majorities. In the event of a tie, the model remains silent regarding the aggregator's behavior. From round 3 onward, all non-aggregators imitate the aggregator, who, for each of them, satisfies the conditions for being an influencer as characterized in Proposition 1.

The naïve model requires a separate analysis for each network. We begin with the Star network, where the non-aggregators (the "leafs") are completely disconnected from one another. In this network, a stable majority is reached in the round in which the aggregator's guess aligns with the majority's guess. This dynamic arises from the behavior of naïve leafs: if a leaf agrees with the aggregator, she maintains her guess; if she disagrees, she may switch. Specifically, if the aggregator's private signal corresponds to the majority signal, then all agents holding that signal continue to report it in all subsequent rounds. However, if the aggregator's private signal corresponds to the minority signal, she switches her guess and triggers possible subsequent switches by the leafs. This dynamics continue until the aggregator's guess aligns with the majority's guess.

In the Connected Spokes network, a single aggregator connects multiple cliques of varying sizes, each containing between three and $\frac{n}{2} - 1$ non-aggregators.[24] Building on Lemma 6 in Section A of the Theoretical Appendix, we show that in any clique with a strict first-round majority, the non-aggregators adopt the majority guess in round 2 and never switch thereafter. In the case of a tie, their guesses remain undetermined until a majority emerges, after which they follow it consistently. The aggregator monitors the overall distribution of guesses across all agents and may continue to switch as long as ties persist within some cliques (for a similar reasoning see Choi et al. (2023)).

Finally, we apply Lemma 6 once again to analyze naïve behavior in the One Gatekeeper network, where the gatekeeper is the sole member of the core clique connected to peripheral hangers-on.[25] We show that, due to the rapid convergence of the core, if the core's majority signal coincides with the global majority signal, most agents guess correctly by round 3 at the latest. However, if the core's majority is incorrect, a substantial share of agents may quickly converge on the wrong guess.

The Single Aggregator networks illustrate the contrasting implications of the two models for information aggregation. In the Bayesian model, agents assume others are myopic Bayesian and recognize that the aggregator's second-round guess reflects all private signals. As a result, they optimally imitate her, and information flows perfectly through the aggregator. In contrast, in the naïve model the aggregator is treated as just another peer, so her influence diminishes with the size of the local neighborhoods. In fact the results show that non-aggregators embedded in cliques often form their final beliefs before even observing the aggregator's guess. Thus, in the naïve model, information rarely flows through the aggregator.

[24] Formally, a Connected Spokes network with aggregator i consists of a collection of cliques $C_1, \ldots, C_m$, each satisfying $3 \leq |C_j| \leq \frac{n}{2} - 1$, such that any two distinct cliques share only the aggregator in common: for all $j_1 \neq j_2$, $C_{j_1} \cap C_{j_2} = i$.

[25] A One Gatekeeper Network is a core periphery network in which the core consists of $n = m + 1$ agents: the aggregator i and the set $C(G) = j_1, \ldots, j_m$. The periphery consists of n agents, $K(G) = k_1, \ldots, k_n$, each of whom is connected only to the aggregator. We assume n is odd.

The Symmetric Core–Periphery Network (Result 6 in Section B of the Theoretical Appendix). The Symmetric Core–Periphery network consists of two equal-sized groups of $\frac{n}{2}$ agents each (n is even and greater than 4): a fully connected core and a periphery of disconnected nodes, each linked to a unique core member. Result 6 in Section B of the Theoretical Appendix characterizes behavior under both the myopic Bayesian and naïve models.

According to the Bayesian model, agents follow their private signals in the first round. If a clear majority emerges in the core (i.e., the difference in signal frequencies exceeds 1), all core members adopt the majority guess from the second round onward, and peripheral agents begin imitating them in the third round. If there is no clear majority, the dynamics depend on whether $\frac{n}{2}$ is even or odd. When $\frac{n}{2}$ is even, information aggregates perfectly: the core agents guess the global majority from the third round, and the periphery imitate them from round 4. When $\frac{n}{2}$ is odd, some core agents are indifferent in round 2, and the model yields no definitive prediction.

Interestingly, under the naïve model, core agents behave identically to their Bayesian counterparts. However, peripheral agents are less predictable: from round 3 onward, they follow the local majority (which may be tied) rather than imitating their connected core member.

These results highlight that, in most cases, core agents —whether naïve or myopically Bayesian — base their decisions solely on the signals within the core. This can be detrimental. Consider the case where $n = 18$: if six core members receive signal w and the remaining 12 agents receive signal b, then all agents converge to the incorrect guess w from round 3 onward, even under the Bayesian model. The reason is that information from the periphery does not reach the core, as each core agent observes only one peripheral neighbor and cannot aggregate across them. In such cases, where aggregation fails despite all agents behaving myopically rationally under common knowledge, we say that information aggregation is impeded by a *structural friction*.[26]

Networks with Two Cores and a Few Bridging Links (Results 7, 8, 9 and 10 in Section B of the Theoretical Appendix). Networks with two cores consist of two internally connected cliques of equal size ($\frac{n}{2}$, with n even and greater than 4), connected by a small number of *bridging links*. In the Two Cores with One Link network, a single bridging link connects agent i from one clique with agent j from the other; these two agents are the *connectors*. In the Two Cores with Three Links network, three agents i_1, i_2, i_3 from one clique are each connected to a common agent j in the other clique (agents i_1, i_2, i_3, and j are the *connectors*).

Under the naïve model, the dynamics follow directly from Lemma 6: from the second round onward, agents follow the majority within their clique (with connectors potentially deviating in round 2). This can lead to aggregation failures when, after the first round, the majority in one clique is W while the majority in the other clique is B, even if a global majority exists.[27]

Bayesian dynamics are more intricate. According to Proposition 1, non-connectors imitate the

[26] Given our experimental parameters ($n = 18$, $q = 0.7$), the probability of incorrect aggregation by myopic Bayesian agents due to structural friction in the symmetric core–periphery network is approximately 2.35%.

[27] Results 8 and 10 assume $\frac{n}{2}$ is odd to avoid tie scenarios for non-connectors. Since we make no assumptions about tie-breaking behavior, such ties may or may not mitigate local failures.

connectors from the third round onward.[28] Successful information aggregation thus hinges on the connectors' ability to communicate across cliques. However, communication becomes difficult when signal distributions induce conflicting local majorities—precisely the cases where the naïve model fails. In such instances, a connector cannot fully convey the information regarding the signals observed within her clique using only a binary action. Nevertheless, for the parameters used in our experiment, we show that when connectors disagree in round 2, successful communication is still possible—as long as agents correctly interpret the information conveyed by *not* switching (see Section B.11 of the Theoretical Appendix for a detailed example).[29] While we demonstrate that successful aggregation is theoretically possible in such cases, it requires an extraordinary high level of reasoning, even under common knowledge of rationality. We refer to these cases—where aggregation is possible in the myopic Bayesian model but requires unusually elevated reasoning—as instances of *surmountable structural friction*.

Finally, Two Cores networks are also susceptible to *structural friction* under Bayesian dynamics. Consider the Two Cores with One Link network with $n = 18$. Suppose that three non-connectors in each core and both connectors receive signal w, while the remaining agents receive signal b. The connectors are indifferent in round 2. If both guess W, they will continue to do so in subsequent rounds due to their agreement—leading to incorrect convergence despite rational behavior.[30]

Table 2 summarizes predicted behavior in all seven networks studied in the experimental lab under both the myopic Bayesian and the naïve models.

4 Aggregate Analysis

This section analyzes the collective performance of subjects in aggregating dispersed private information across different network structures. We begin by describing the data (Section 4.1) and introducing an *aggregate learning index* used to quantify information aggregation (Section 4.2). We then use this index to examine how efficiently each network structure facilitates information transmission as a function of the aggregate quality of the private signals and use it to evaluate the performance of the Bayesian model (Section 4.3). Next, we assess the extent to which observed aggregation failures can be attributed to the structural frictions defined earlier (Section 4.4). We then turn to evaluate the performance of the naïve model in predicting aggregate outcome (Section 4.5). Finally, we examine the dynamics of the aggregation process, focusing on how long it takes for learning to converge (Section 4.6).

[28] Proposition 1 does not apply to non-connectors in the clique containing i_1, i_2, i_3 in the Two Cores with Three Links network. As shown in Result 9.2, if i_1, i_2, and i_3 are unanimous in period $t-1$, then non-connectors follow them in period t; otherwise, they infer indifference and guess randomly.

[29] Similar dynamics of information exchange appear in Geanakoplos and Polemarchakis (1982). However, in our setting, communication is through binary actions rather than posteriors. Thus, the general results in Geanakoplos and Polemarchakis (1982) do not directly apply. The proofs of Results 7 and 9 are specific to our experimental parameters.

[30] The probability of such a signal distribution is approximately 0.69%. Since we make no assumptions about tie-breaking behavior, this is an upper bound on the likelihood of this type of structural friction. The corresponding probability in the Two Cores with Three Links network is 0.98%.

Network	Myopic Bayesian Model	Naïve Model
Complete	Always correct from $t \geq 2$ (1).	Always correct from $t \geq 2$ (1).
Star	Always correct from $t \geq 3$ (2).	If the aggregator recieves the majority's signal, most are correct from $t \geq 2$. Otherwise, indeterminate (3).
Connected Spokes	Always correct from $t \geq 3$ (2).	Non-aggregators guess by local majority from $t \geq 2$. The aggregator aggregates their choices (4).
One Gatekeeper	Always correct from $t \geq 3$ (2).	Non-aggregators in the core choose by local majority from $t \geq 2$. In most cases, the others follow whether correct or incorrect (5).
Symmetric Core Periphery	In most cases, the core members aggregate only their own signals. The leafs imitate from $t \geq 3$ (6). *Possible Structural Frictions*	In most cases, the core members aggregate only their own signals. Most leafs are correct (6).
Two Cores with One Link and Two Cores with Three Links	If both cores agree, they follow the agreed guess from $t \geq 2$. *Possible Structural Frictions* Otherwise, slow convergence to the correct guess (7, 9). *Possible Surmountable Structural Frictions*	Agents guess by the majority of their local core from $t \geq 2$ (8, 10).

Table 2: Summary of Theoretical Predictions

Notes: Formal statements and proofs appear in Section B of the Theoretical Appendix and referenced throughout the table in the parenthesis.

4.1 Data

We collected data from 410 games played across the seven network structures described in Section 2.1 (see Table 1 for details). Our analysis follows three guiding principles. First, in the 10 games where the majority of private signals did not match the true state selected by the computer, we redefine the state to align with the majority of signals (see Footnote 13 for discussion). Second, we exclude games in which the number of each signal type is equal, since in such cases every guess is considered correct by definition. Third, we exclude games where participants failed to converge—defined as cases in which at least one participant continued to revise their guess beyond round 50 (see Footnote 8). These two exclusions eliminate 48 of the 410 games (11.7%).[31]

We refer to signals that match the majority of all signals in the network as correct, and those that do not as incorrect (or wrong). To capture the informativeness of the initial signal distribution at the network level, we categorize each game into one of three signal-quality levels: weak, average, or strong. Specifically, we label games with *weak signals* as those in which 10 or 11 participants receive correct signals; games with *average signals* as those in which 12 or 13 participants receive

[31] The results are robust to the inclusion of games in which the group reached round 50 (available from the authors upon request).

correct signals; and games with *strong signals* as those in which at least 14 participants receive correct signals. This classification ensures a reasonably balanced distribution across signal categories: 22% of games begin with weak signals, 40% with average signals, and 38% with strong signals.

4.2 Defining a Measure of Information Aggregation

We use the term learning to refer to changes in participants' posterior beliefs resulting from observing others' actions. In our design, we observe participants' coarse actions rather than their beliefs, so learning is not always directly observable. Indeed, we can only definitively state that a participant has learned if they report a guess differing from their initial signal. We say a participant learns correctly if they initially receive an incorrect signal but ultimately make a correct guess. In contrast, incorrect learning occurs when a participant initially receives a correct signal but ultimately reports an incorrect guess. To evaluate how network structure affects information aggregation, we construct an **aggregate learning index** (ALI), which captures these observable instances of learning.[32]

Definition. For each game g, let CS^g denote the number of correct signals and IS^g denote the number of incorrect signals (CS^g+IS^g equals the network size in game g). For every round t in game g, let CG_t^g denote the number of correct guesses submitted. Then, we define:

$$\text{ALI}_t^g = \begin{cases} \frac{CG_t^g - CS^g}{IS^g} & CG_t^g > CS^g \\ 0 & CG_t^g = CS^g \\ \frac{CG_t^g - CS^g}{CS^g} & CG_t^g < CS^g \end{cases}$$

The Aggregate Learning Index (ALI) is an intuitive measure of overall information aggregation success in game g at the end of round t. Specifically, if the number of correct guesses in round t exceeds the number of correct initial signals, ALI represents the net fraction of participants who learned correctly relative to the total number of participants with incorrect initial signals. Conversely, if the number of correct guesses falls below the number of correct initial signals, ALI is negative and reflects the net fraction of incorrect learners relative to the total number of participants with initially correct signals.

Importantly, ALI takes values in the interval $[-1, 1]$. A value of 1 represents absolute information aggregation: all participants with incorrect signals revise correctly, and those with correct signals retain their initial signals. Conversely, a value of -1 indicates complete aggregation failure: all participants with correct signals revise incorrectly, and no participant with incorrect signals updates. When no participant deviates from their initial signal, ALI equals zero.

A notable property of ALI is its composition invariance. Consider two games, g and h, each with 18 participants. In both games, 12 participants initially receive correct signals ($CS^g = CS^h = 12$) and 6 receive incorrect signals ($IS^g = IS^h = 6$). Suppose further that in round t, 14 correct guesses occur in both games ($CG_t^g = CG_t^h = 14$). In game g, among the 14 correct guesses, 12 participants

[32] See Choi et al. (2023) for a similar index. Choi et al. (2005) introduce the "stability measure": a related but distinct index based on switching behavior.

initially received correct signals, whereas in game h, only 8 initially received correct signals. Despite these differences, in both games, the ALI equals $\frac{1}{3}$. Thus, ALI provides a high-level measure of learning outcomes, abstracting from detailed individual-level learning outcomes. One might argue, however, that the extent of learning in game h is greater than in game g, because a higher fraction of participants with initially incorrect signals learned correctly, even after accounting for participants with initially correct signals who guessed incorrectly. To capture these individual-level differences, we introduce a second measure, the **individual learning index** (ILI), which measures the success of learning at the individual level. We define and discuss the ILI in Section C.1 of the Empirical Appendix, provide the cumulative densities of ALI and ILI in Section C.2 and demonstrate that our results remain robust when replacing ALI with ILI in Section C.3.

4.3 The Performance of the Bayesian Model

Figure 2 presents scatter plots of network-specific end-game ALIs as a function of the fraction of correct initial signals for each network. The bubble size corresponds to the number of observations with identical outcomes. In discussing empirical patterns exhibited in Figure 2, we focus on absolute success and failure statistics in learning and on the relationship between the extent of learning and the initial signal distribution.

We use the Complete network as the benchmark for our analysis, as it imposes no restrictions on nodes' connectivity and thus offers the greatest potential for aggregating dispersed private information about the state of the world. Information aggregation is not absolute even in the Complete network, but it features two important properties. First, participants in the Complete network almost always learn correctly, in aggregate.[33] Second, the rate of aggregate correct learning responds positively and monotonically to the quality of initial signals (i.e. games with stronger signals achieve higher end-game ALIs).

Figure 2 provides a natural partition of the non-complete networks we studied. The first group consists of networks in which ALI responds positively and monotonically to signal quality, while in the second group, no such association is found. This difference maps clearly onto the two structural features discussed in Section 2.1. The first group—networks where ALI and informativeness positively correlate—comprises the Symmetric Core Periphery network, the Two Cores with One Link network, and the Two Cores with Three Links network. Each of these networks features one or two cliques of 9 nodes; we label them *Cluster(s)* networks. The second group—networks where ALI and informativeness do not positively correlate—comprises the Star network, the Connected Spokes network, and the One Gatekeeper network. Each has a single node connected to all others; we label these *Single Aggregator* networks. Although the One Gatekeeper network features both a central node and a size-9 clique, it is clearly insensitive to the distribution of initial signals, behaving like the other *Single Aggregator* networks rather than the *Cluster(s)* networks.

[33] We observed only one game (out of 46) in which most participants converged to the minority signal (the only dot below the zero horizontal line in the Complete network diagram in Figure 2). The initial distribution of signals in that case was 10-8.

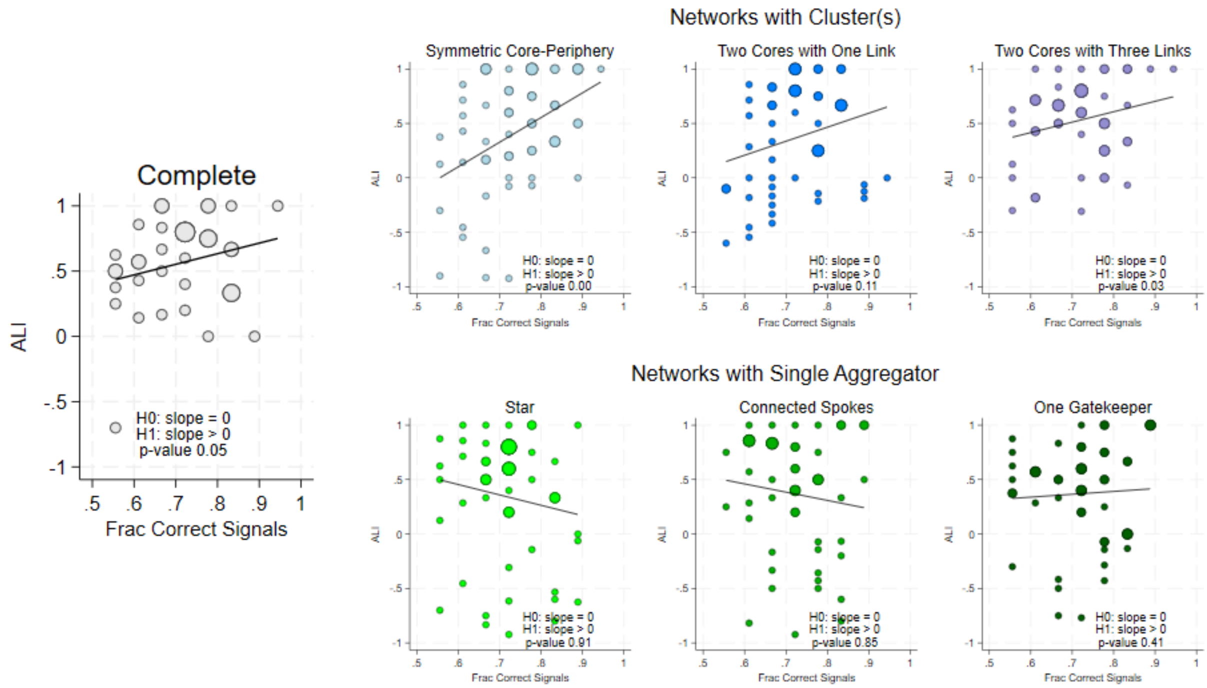


Figure 2: Aggregate Learning Indices, by network

Notes: The fraction of correct signals is on the horizontal axes. The final round ALI is on the vertical axes. The size of the bubble corresponds to the number of observations. The straight lines are the linear fit. The p-value reports the test for the null hypothesis that the slope of the linear fit equals zero against a one-sided, positive, alternative in a regression using clustered standard errors at the session level.

Given the stark differences between the two groups, and consistent with our focus on network architectural features' effects on information aggregation, much of the analysis below pools data from networks within each group. We compare the performance of *Cluster(s)* and *Single Aggregator* groups against each other and against our benchmark, the Complete network.

Figure 2 also shows that absolute aggregation success ($ALI = 1$), while predicted by the Bayesian model in most cases, is rarely achieved in practice. Moreover, aggregation failures are fairly common. We define a *complete failure of information aggregation* as a case in which a majority of participants make incorrect guesses in the final round. A *relative failure of information aggregation* occurs when the final-round ALI is negative. Relative failure is a necessary but not sufficient condition for complete failure. Panel A in Figure 3 reports the frequency of both types of failures, as well as absolute success, for each network.[34] Consistent with our observations above, panel A highlights two patterns. First, while relative or complete information aggregation failures occur almost never in the

[34] Mueller-Frank and Neri (2015) report an absolute aggregation rate of 70% for a complete network of size 5, and 46.5% for a star network of the same size. Grimm and Mengel (2020) find 27% absolute aggregation in networks of size 7 that the myopic Bayesian model predicts will reach consensus on the correct state. Figure 3c in Choi et al. (2023) shows a 26% relative failure rate and no instance of absolute aggregation in networks of size 40.

Panel A: Failure Rates and Absolute Aggregation

Network	Relative Failure	Complete Failure	Absolute Aggregation	
Complete	3%	3%	16%	●
Single Aggregator				
Star	23%	18%	11%	◆ (grey)
Connected Spokes	27%	15%	15%	◇ (black)
One Gatekeeper	22%	12%	12%	◇ (grey)
Cluster(s)				
Sym Core Periphery	17%	12%	24%	□ (black)
Two Cores One Link	29%	9%	15%	□ (grey)
Two Cores Three links	10%	2%	17%	■ (grey)

Panel B: Final Guesses as a function of Signals' Quality

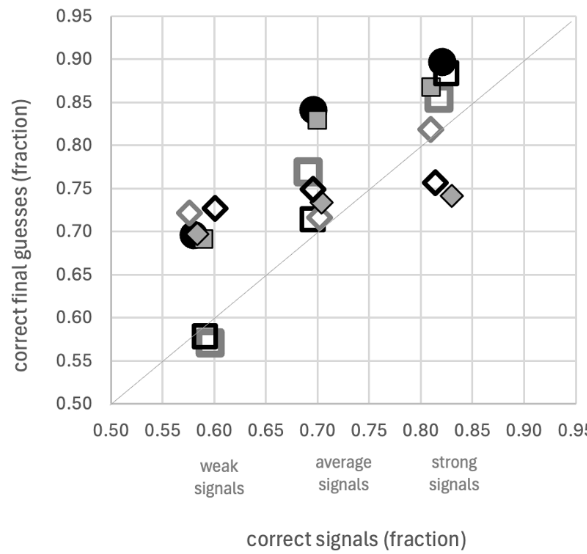


Figure 3: Rates of Information Aggregate Successes and Failures, by Network

Notes: Panel A reports the frequency of relative information aggregation failures ($ALI < 0$) in the second column; the frequency of complete aggregation failures (the final majority guess is incorrect) in the third column; and the frequency of absolute aggregation success ($ALI = 1$) in the fourth column. Panel B presents, on average by network structure, a scatter plot showing the share of correct final guesses as a function of the percentage of correct private signals, grouped by signal quality: weak, average, and strong (as defined in Section 4.1). The legend is placed in the right-most column of Panel A.

Complete network, these failures do occur in significant proportions in both *Single Aggregator* and *Cluster(s)* networks. Second, absolute aggregation of information is rare in all networks, including the Complete one. Panel B in Figure 3 presents a scatter plot showing the average share of correct final guesses as a function of the percentage of correct private signals, grouped by network structure and signal quality. Panel B of Figure 3 and Figure 2 further illustrate that *Cluster(s)* networks frequently fail when initial signals are weak, whereas *Single-Aggregator* networks perform relatively poorly in cases with strong initial signals, compared to other network structures.

Before we explore the performance of Naïve model, we show that structural frictions cannot account for the observed information aggregation failures.

4.4 Structural Frictions

Recall from Section 3 that we define a structural friction as a situation in which information aggregation fails even under common knowledge that all agents are myopically Bayesian. We also introduced the notion of a surmountable structural friction: a case in which aggregation is theoretically possible under these same assumptions but requires unusually sophisticated reasoning.[35]

[35]By Table 3 in Grimm and Mengel (2020), they study one case of structural friction (Kite 1) and two cases of surmountable structural frictions (Circle 2 and Kite 2).

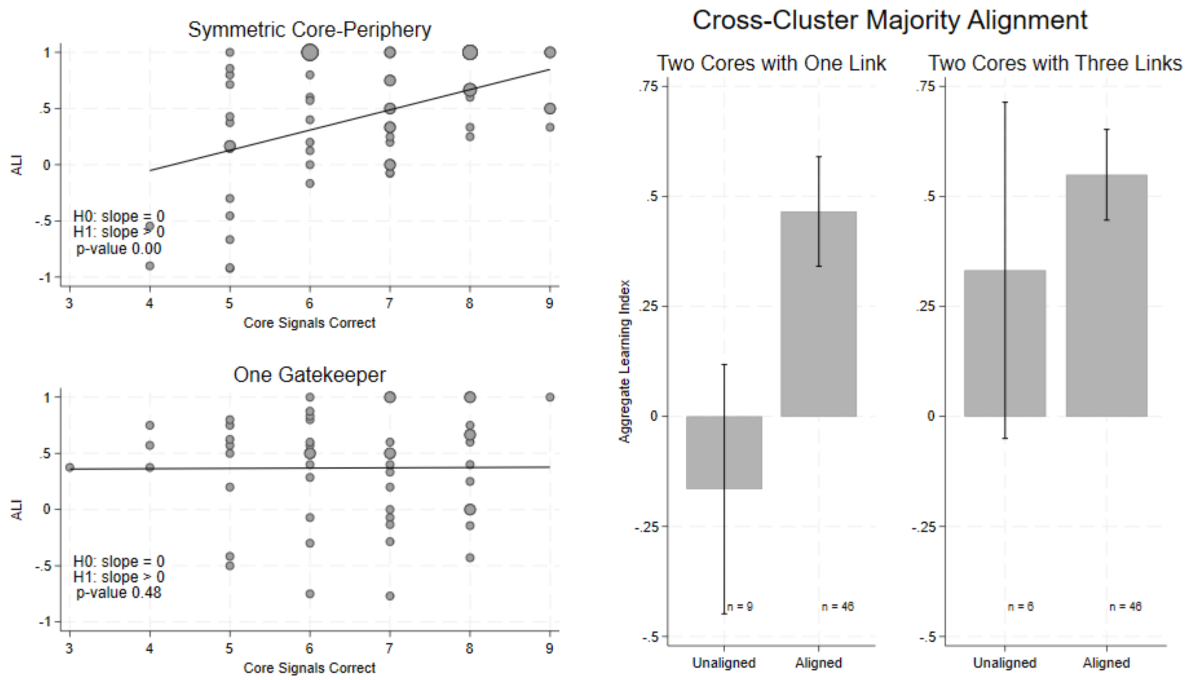


Figure 4: Structural Failures

Notes: The figures on the left plot end-game ALIs in the Symmetric Core Periphery and One Gatekeeper networks as a function of the number of correct signals in the core. The figures on the right plot end-game ALIs in the Two Cores with One and Three Links networks as a function of the alignment of majority signals across the two cliques.

The results in Section 3.4 show that in *Cluster(s)* networks, certain signal distributions give rise to both types of frictions. We next examine the extent to which these structural frictions account for the aggregation failures documented in Figures 2 and 3.

One Gatekeeper vs. Symmetric Core Periphery Both networks feature a completely connected core and a periphery in which each node connects to a single core member. The only structural difference lies in the periphery's pattern of connections: in the Symmetric Core Periphery network, each peripheral node connects to a different core member; in the One Gatekeeper network, all peripheral nodes connect to the same core member—the "Gatekeeper." Section 3.4 predicts no structural failures for the One Gatekeeper network, but identifies potential frictions in the Symmetric Core Periphery network when the core's majority is either narrow or incorrect. Panel A of Figure 4 shows that average ALI in the One Gatekeeper network is largely insensitive to the signal distribution in the core. In contrast, average ALI in the Symmetric Core Periphery network increases with the number of correct signals in the core—as predicted by structural considerations.

The Two Cores Networks In the Two Cores networks, Section 3.4 predicts that aggregation is straightforward when both cliques have the same majority signal, although errors can still occur with low probability. However, when the majority signals in the two cliques conflict, aggregation becomes extremely difficult: success may require multiple iterations of complex inference by the connectors, without switching. Panel B of Figure 4 confirms this prediction—especially for the Two Cores with One Link network—showing that aggregation often fails when the two cliques' majority signals are misaligned.

Not the Whole Story The cases discussed above demonstrate that some network structures are prone to aggregation difficulties for specific signal distributions due to structural frictions. This is the case for weak and average signals in some of the *Cluster(s)* networks (panel B in Figure 3). However, the myopic Bayesian model does not predict any such difficulties in the *Complete* or *Single Aggregator* networks: these networks are expected to aggregate information fast regardless of the signal distribution (Section 3.4). However, Figure 2 and Panel B in Figure 3 reveal systematic deviations from this prediction: *Single Aggregator* networks often fail to aggregate information even when initial signals are average or strong. We, therefore, conclude that while structural frictions explain some failures, they cannot fully account for the observed performance. We propose that the remaining failures stem from *behavioral frictions* that interact with network architecture and signal distributions. In the remainder of the paper, we explore these behavioral frictions.

4.5 The Performance of the Naïve Model

As discussed in Section 3.4, the network-level predictions of the naïve model differ substantially from those of the myopic Bayesian model. In particular, the naïve model rarely predicts absolute aggregation. This is largely because leafs and small cohesive groups often struggle to learn the true state under naïve updating. An exception is the Complete network, where both models predict absolute aggregation, yet subjects' actual performance falls short of this benchmark (see Section 4.3).

Because the naïve model often yields indeterminate predictions, we test a key feature of its behavior: that clique members with limited external connections should converge early and never revise their guess. This prediction, formalized in Lemma 6 (Section A of the Theoretical Appendix), states that such agents should adopt the round-1 local majority from period 2 onward and stick to it.

To test this implication in its most obvious form, we focus on clique members with no links outside their clique. In the networks we study, these agents should behave according to Lemma 6, using the local majority of round-1 guesses to determine their actions in all subsequent rounds. Specifically, we examine four groups of such positions: (i) non-aggregators in the Connected Spokes network, (ii) non-aggregator clique members in the One Gatekeeper network, (iii) non-connectors in the Two Cores with One Link network, and (iv) non-connectors in the Two Cores with Three Links network. For each case (excluding ties), we compute the local majority after round 1 and check whether the subject followed it consistently thereafter.

In total, we identify 2,769 relevant instances. In 68.3% of them, subjects adhered to Lemma 6.

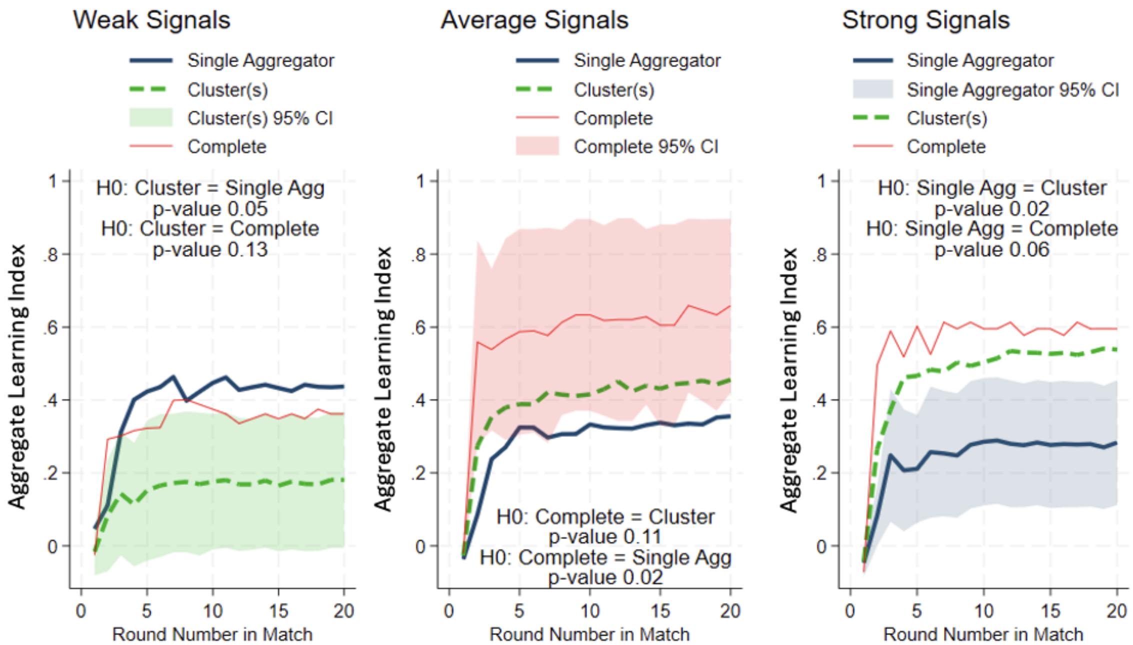


Figure 5: Evolution of ALI as the Game Progresses

Notes: The figure presents the average ALI per round, across network groups and signal quality. For readability, we present the 95% confidence intervals only for the network group that most differs from the others, using clustered standard errors at the session level. The reported p-values use these clustered standard errors to evaluate the null hypotheses that the most different group's mean ALI differs from the other groups' ALIs in round 20. As there are no noticeable movements beyond round 20, the horizontal lines end there. See Section C.5 of the Empirical Appendix for robustness analysis.

Adherence was highest among non-connectors in the Two Cores with One Link (75.6%) network, and lower among non-connectors in the Two Cores with Three Links (68.3%) network, non-aggregators in the Connected Spokes (63.9%) and One Gatekeeper (60.5%) networks. Given that the initial signal is correct in approximately 70% of cases, these adherence rates are modest, at best. This means that these participants exhibit very low rates of local information aggregation.

Taken together—the fact that a significant fraction of subjects violate a clear implication of the naïve model, and the surprisingly poor performance in the Complete network—we conclude that relying solely on local neighborhood information does not provide an adequate explanation for the aggregate behavior observed in the experiment.

4.6 Dynamics

Figure 5 shows how ALI evolves over rounds, separately for games with weak, average, and strong initial signals. When signals are not weak, the *Complete* network reaches an ALI of around 0.5 as early as round 2. When signals are weak, it still achieves an ALI of approximately 0.3 by round 2. In both cases, performance improves slightly in subsequent rounds. The *Cluster(s)* networks perform poorly when signals are weak but improve markedly with signal quality, eventually matching the *Complete* network when signals are strong. By contrast, the *Single Aggregator* networks perform similarly to the *Complete* network when signals are weak but fail to improve as signal quality increases.[36]

Another key insight from Figure 5 is that the first three rounds largely determine the aggregate outcome. ALI stabilizes early, with minimal change after round 3 (see Grimm and Mengel (2020) and Choi et al. (2023) for similar observations). This motivates us to focus the positional-level analysis on behavior during the first three rounds.[37]

4.7 Summary

First and foremost, our analysis establishes that *network structure has a profound effect on long-run outcomes*, and that this effect depends critically on the quality of initial information. In particular:

Finding 1. *The Complete network aggregates information better than all other networks. However, it does not perform as perfectly as both theories predict.*

Finding 2. *Single Aggregator networks perform on par with the Complete network when initial information is poor, but fail to improve as information quality increases. As a result, they frequently fail to aggregate information even when the vast majority of signals are correct. Single Aggregator networks perform significantly worse than predicted by theory.*

Finding 3. *Cluster(s) networks respond positively to signal quality and match the Complete network's performance when initial signals are strong. However, when signals are weak, they often fail to aggregate information—consistent with the theory-based notions of surmountable and insurmountable structural frictions.*

Finding 4. *Relying solely on local neighborhood information, as advocated by the naïve model, does not provide an adequate explanation for the observed aggregate behavior.*

Finding 5. *Behavior in the first three rounds is decisive for determining network outcomes.*

5 Positional Analysis

In this section, we study individual behavior by position in the network. Following Finding 5 we focus on the first three rounds. Our benchmarks are the behaviors predicted by the Bayesian model

[36] Section C.6 of the Empirical Appendix evaluates a myriad of regression specifications to test for differences between the final round learning index value between the different types of network. The results in most specifications comparing the Complete network with Cluster(s) networks and the Complete network with Single Aggregator networks demonstrate a statistically significant contrast. The contrasts between the Single Aggregator and the Cluster(s) networks are less clearly significant, especially with substantial controls absorbing the variation in learning index values.

[37] Chandrasekhar et al. (2020) also restrict their individual-level analysis to the first three periods. Their justification differs from ours: "from period 4 onward, most networks enter a zero-probability information set—that is, at least one agent observes behavior that cannot be reconciled with either Bayesian or Naïve reasoning" (p. 23).

		Single Aggregator networks			Cluster(s) networks		
	Complete	*Star*	*Connected Spokes*	*One Gatekeeper*	*Core Periphery*	*Two Cores One Link*	*Two Cores Three Links*
All nodes	92%	93%	92%	91%	94%	92%	91%
Aggregators		91%	94%	94%			
Cluster members	92%		92%	90%	94%	92%	92%
Leafs		93%		92%	95%		
Connectors						90%	88%

Table 3: First-round guesses, by network and position

Notes: Frequency of "correct" first-round guess is reported, where correct indicates a guess coinciding with one's private signal. Aggregators are the unique nodes in the network that are connected to all other nodes. Cluster members are members of a clique of size at least 3 that are not connected to nodes outside the clique. Leafs are nodes with a single link. Connectors are the nodes in the two cores networks that maintain cross-clique links (see Table 9 in Section D.1 of the Empirical Appendix for a detailed classification).

and by the naïve model as described in Section 3.

Throughout this section, we employ regression analysis to examine heterogeneity in participants' behavior as a function of their local information environment. To account for differences in aggregate information distributions and network structures, we include session-game fixed effects. Standard errors are clustered at the individual level to account for within-subject correlation across games.

5.1 First Round Guesses

Both the Bayesian model and the naïve model predict that the first guess should reveal one's private signal since it is correct with probability 70% conditional on the state (Lemma 1 and Definition 1). Moreover, we believe that reporting one's own signal indicates a basic understanding of the game and its reward scheme. First-round guesses match their signals in 92.2% of the cases.[38] Table 3 reports these rates by network and by position. Importantly, there is little variation in the tendency to report one's own signal in the first round of a game across network structures and network positions.[39]

Across ten games within a session, 72.5% of participants report their signal as their first guess in all ten games, and 91.2% misreport at most twice. More than 40% of the misreports (210 out of 510) were made by subjects that misreport at most twice. This suggests that most misreports reflect random "trembling hand" errors rather than strategic considerations or misunderstanding of the game.[40] At the group level, in 94.2% of games with an initial signal imbalance that converged,

[38] Choi et al. (2005) report a first-round mistake rate of 5.8% in networks of size 3. Mueller-Frank and Neri (2015) report rates of 3% and 6% in their 2-urn treatments with 5 and 7 agents, respectively. Jiang et al. (2023) report a first-round mistake rate of 1.7% in networks of size 7.

[39] The Symmetric Core Periphery Network sessions had participants' first-round reports match their signals in 94% of cases, a small difference that proves to be statistically significant even after clustering and multiple-comparison adjustments. None of the other network types showed pairwise differences that were statistically significant. Further, there were no statistically significant differences in the frequency with which the first guess matched a subject's signal by its position in the network. See Section D.1 of the Empirical Appendix for further details.

[40] Almost half (45.2%) of the subjects that misreported at least three times in the first round are classified as probability matchers (see Section A.4 of the Empirical Appendix).

	Benchmark		*Single Aggregator* networks						*Cluster(s)* networks					
	Complete		Star		Connected Spokes		One Gatekeeper		Symmetric Core Periphery		Two Cores One Link		Two Cores Three Links	
Overall	87%		79%		88%		77%		87%		84%		82%	
	maj	min	maj	min	maj	min	maj	min	maj	min	maj	min	maj	min
All nodes	98%	62%	97%	48%	95%	54%	96%	35%	95%	58%	95%	57%	94%	54%
Single Aggregators			97%	48%	91%	61%	100%	59%						
Cluster members					95%	53%	95%	31%	95%	58%	95%	56%	95%	52%
Connectors											97%	71%	90%	65%

Table 4: Second-round guesses, by network and position

Notes: The average frequency of "correct" guesses is reported for all nodes with two or more local friends. Columns "maj" and "min" refer to cases in which a participant's first round guess is part of round 1 local majority or minority, respectively. The round 2 guess is considered "correct" if it matches the local round 1 majority taking into account the participant's own round 1 guess. Leafs are excluded, as they err only if guessing against their signal when their neighbor's guess matches it. In addition, we exclude local ties, where a tie is also defined relative to one's first round guess. Section D.2.1 of the Empirical Appendix reports similar results under different definitions of the "correct" round 2 guess and minority status.

misreports did not alter the majority signal.[41]

Finding 6. *Subjects tend to report their private signals in the first round of the game. Mistakes are relatively rare and are not systematic across network structures or network positions.*

5.2 Second Round Guesses

As we saw, most subjects truthfully report their private signal in the first round of a game, and those who do not have no particular bias. Hence, the predicted second-round behavior, both for Bayesian agents and for naïve agents, entails reporting the majority of first-round guesses one observes in her local neighborhood augmented by their own signal (Lemma 3 and Definition 1). If there is an equal number of guesses of each color, then the subject should be indifferent.[42]

Table 4 documents the frequency with which subjects guess "correctly" in round 2 conditional on whether their first-round guess aligns with the majority of first-round guesses in their local neighborhood. This analysis focuses on nodes with at least two local neighbors (i.e., excluding leafs) and omits cases where local ties occur. Subjects whose first-round guess agrees with the local majority in the first round make a correct guess in round 2 almost always (90% or more), regardless of their network position. In contrast, when their first-round guess contradicts the first-round local majority, the probability of a correct guess in round 2 varies substantially across network structures, ranging from 31% for cluster members in the One Gatekeeper network to 71% for connectors in the

[41] The 21 cases in which first-round misreports reverse the majority signal account for only 18.6% of the relative failures and 23.7% of the complete failures of information aggregation documented in Panel A of Figure 3.

[42] Our theoretical analysis does not impose a tie-breaking rule. In the experiment, ties occurred in 1,287 second-round decisions (17.4%), with about 75% involving leaf nodes. In 75.37% of tie cases, subjects repeated their first-round guess. This fraction is similar if we instead measure tie-breaking relative to the subject's initial signal rather than their first-round guess (74.27%).

Two Cores with One Link network.[43]

First Behavioral Friction: Under-Reaction to New Information. The Bayesian model predicts that in round 2, all agents should optimally switch their guess to the majority guess from round 1. Furthermore, under the naïve model, subjects are expected to adopt the majority guess from round $t-1$ as their guess in round t ($t > 1$). However, the behavior observed in Table 4 deviates from the predictions of both the Bayesian and naïve models. Instead, the behavior appears to be consistent with the well-documented behavioral bias known as *under-reaction to new information*. In our setting, the new information consists of the first-round guesses of a subject's direct neighbors, which become observable only at the end of round 1. The under-reaction bias manifests when subjects recognize that the majority of their direct neighbors received a signal different from their own—based on their round 1 reports—but nonetheless fail to switch to the majority signal as frequently as would be expected.[44]

This behavior is hardly surprising as under-reaction to new information is one of the more stable and well-documented empirical deviations from Bayesian predictions.[45] Conlon et al. (2022) find that subjects who exert effort to uncover information overweight their private signals relative to their partner's, which they interpret as an ownership effect. Esponda et al. (2023) show that subjects overweight private signals relative to group-level information. Augenblick et al. (2025) use the cognitive imprecision model of Woodford (2020) and find under-reaction with precise signals and over-reaction with weak signals. Ba et al. (2024, 2025) combine noisy cognition and representativeness, predicting under-reaction when the state space is simple, signals precise, and priors flat, and over-reaction when the environment is more complex, signals noisier, and priors more concentrated. The environment in our experiment aligns with conditions predicted to generate under-reaction in both Augenblick et al. (2025) and Ba et al. (2024, 2025) models.

Table 5 uses a linear probability model to analyze under-reaction in round 2 as a function of local environment and network position. The dependent variable equals 1 if the round 2 guess matches the local round 1 majority; the key regressor indicates whether the individual was in the round 1 minority. Additional controls capture the extent of local consensus in round 1, network position, and individual characteristics.

The first regression shows that being in the local majority in round 1 is associated with a 94.8% probability of making a correct guess in round 2. Belonging to the minority reduces this probability by approximately 40 percentage points, across networks and positions. The other regressions add

[43] Choi et al. (2005) study networks of size 3 with signal accuracy $\frac{2}{3}$. In their full information treatment, subjects in the complete network were incorrect in round 2 in 13% of cases, while aggregators in the star network were incorrect in 11.1% of cases. Choi et al. (2023) study networks of size 40 with signal accuracy 0.7 and report that approximately 20% of subjects switched their guesses between the first and second rounds. In addition, their Table EC.4 reveals that 10%-12% of the subjects were incorrect in at least one of the first two rounds. Neither study considers whether a subject was in the majority or minority at the end of round 1.

[44] An extreme form of Under-Reaction to New Information is to adhere to one's initial private signal throughout the game—a behavior often labeled "stubbornness" in the social learning literature. Choi et al. (2023) report in their supplementary material that 25%–30% of their subjects exhibit such behavior. In our data, the rate of stubbornness ranges from 3% to 16%.

[45] See the surveys by Benjamin (2019), Enke (2024) and Section 6.2.1 in Bikhchandani et al. (2024).

	Dependent Variable: Correct Round 2 Guess All Non-Leaf Nodes			
	Baseline Model	First Order Model	Interaction Model	Interactions with Controls
Constant	0.948***	0.953***	0.942***	0.923***
	(0.00499)	(0.0269)	(0.0274)	(0.0322)
Minority Characteristics				
In R1 Minority	-0.396***	-0.414***	-0.358***	-0.346***
	(0.0190)	(0.0200)	(0.0475)	(0.0468)
Local Minority Size		-0.175***	-0.0416	-0.0331
		(0.0495)	(0.0465)	(0.0459)
In R1 Minority × Local Minority Size			-0.820***	-0.819***
			(0.127)	(0.126)
Node Characteristics				
Node Degree Centrality		0.0935*	0.0401	0.0310
		(0.0488)	(0.0498)	(0.0493)
Node Degree Centrality × In R1 Minority			0.278***	0.288***
			(0.0762)	(0.0747)
Individual Controls				
Incorrect Round 1 Guess				-0.120***
				(0.0305)
Gender				0.00885
				(0.0123)
Probability Matching				-0.0493***
				(0.0149)
Risk Aversion				0.0505**
				(0.0255)
R-squared	0.243	0.245	0.262	0.275
# of Observations	4,310	4,310	4,310	4,310
# of Clusters	756	756	756	756
# of Session-Game Fixed Effects	359	359	359	359

* p<0.05, ** p<0.01, *** p<0.001. Standard errors in parentheses

Table 5: Determinants of Second-Round Guesses

Notes: These are linear regressions with clustering at the participant level including session-game fixed effects. The sample includes only nodes with two or more neighbores and excludes local ties. In R1 Minority is an indicator that equals one when R1 guess was not the most popular in one's local neighborhood in the first round. Local Minority Size is the percentage of the local minority in the neighborhood. Node degree centrality is calculated as the number of neighbors divided by the largest number of neighbors one can have in our networks (17). Individual controls include the risk attitude measure, the probability matching measure, the indicator of submitting a wrong guess in the first round, and gender. Section D.2.2 of the Empirical Appendix presents robustness checks for different regression model specifications.

two insights. First, for minority members, minority size is negatively correlated with correctness: large minorities double the adverse effect (minority size ranges from 0 to $\frac{4}{9}$). Second, the negative impact of being in the minority is partially mitigated by large neighborhood size. That is, more connections help minority members better incorporate new information.

Finding 7. *Across networks and positions, participants imperfectly aggregate local information in round 2, systematically under-reacting to neighbors' first-round guesses when in the local minority. The extent of under-reaction depends on the strength of the observed evidence, determined by neighborhood size and majority-minority composition of their local neighborhood.*

	Single Aggregator networks						*Cluster(s)* networks					
	Star		Connected Spokes		One Gatekeeper		Symmetric Core Periphery		Two Cores One Link		Two Cores Three Links	
	same	diff	same	diff	same	diff	same	diff	same	diff	same	diff
Leafs	94%	46%			96%	43%	97%	60%				
Cluster members			97%	36%	95%	29%			97%	21%	96%	25%

Table 6: Third-round imitation frequencies, by position and agreement with the influencer in R2.

Notes: We report how often a leaf's or cluster member's round 3 guess matches their influential friend's round 2 guess in cases where imitation is optimal. We distinguish between cases where their own round 2 guess agrees with the influential neighbor's (column "same") and where it differs (column "diff").

5.3 Third Round Guesses

The third round is the first stage at which subjects can incorporate information from network members to whom they are not directly connected. In the naïve model, this occurs mechanically: third-round guesses reflect the second-round guesses of direct neighbors, which were themselves shaped by the first-round guesses of more distant agents. In the Bayesian model, players should begin to exploit the network structure through sophisticated inference to refine their guesses. In practice, as shown in Finding 5, most information aggregation is completed by the end of the third round, making it a particularly important stage to analyze.

In Section 3.2, we introduced the heuristic termed *imitation*: agent i imitates agent j if $\forall t > 2 : a_i^t = a_j^{t-1}$. Proposition 1 identifies network positions where Bayesian agents should optimally imitate a neighbor—specifically, the unique neighbor j who is strictly better informed than agent i and all i's other neighbors (agent k is better informed than agent l if $B(l) \cup \{l\} \subset B(k) \cup \{k\}$). Section 3.4 applies this result to the networks we study. In *Single Aggregator* networks, all non-aggregators should imitate the aggregator (Result 2). In the Symmetric Core Periphery network, each leaf should imitate its core neighbor (Result 6). In the Two Cores with One Link network, non-connectors should imitate their connector (Result 7). Finally, in the Two Cores with Three Links network, non-connectors in the core with a single connector should imitate that connector, and those in the core with three connectors should imitate whenever the connectors' prior-round guesses are unanimous (Result 9). We refer to the agent who should be imitated in these cases as the *Influencer*.

Table 6 reports how often subjects' behavior aligns with imitation of the influencer in cases where Bayesian agents should optimally imitate. When players agreed with the influencer in round 2, they typically maintained the same guess in round 3 (94%–97% across networks). However, when subjects disagreed with the influencer in the second round, they frequently persisted with their round 2 guess rather than switching. Notably, players with smaller local networks (especially leafs) were more likely to imitate the influencer following disagreement.

Second Behavioral Friction: Under-Imitation. Table 6 shows that subjects frequently fail to optimally imitate their influential neighbors following disagreement in round 2—a pattern we refer to as *under-imitation*. This behavior is consistent with findings from sequential social learning experiments, where subjects tend to under-imitate predecessors when doing so requires acting against their private signal (Weizsäcker (2010); Ziegelmeyer et al. (2013)).

Although both under-imitation and under-reaction to new information involve a failure to switch, they are conceptually distinct. Imitation requires a more sophisticated understanding of the network: to decide whether to imitate a neighbor, a subject must consider not only the neighbor's action but also their position in the network, including the connectivity of their neighbors. In contrast, reacting to new information depends solely on the agent's immediate environment. The two behaviors also differ in cognitive demands: imitation involves mechanically copying a neighbor's previous guess (in every period), while responding to new information typically requires a one time computation, such as counting. More broadly, learning can be motivated either by a desire to access others' private signals or by the belief that someone else is better equipped to interpret the environment. Under-reaction reflects a failure of the former—insufficient use of others' private information—while under-imitation reflects a failure of the latter, namely, the inability to recognize that a neighbor may be better informed about the state of the world (see Amelio (2024)). Finally, Table 6 and Finding 7 highlight a key empirical difference between these two frictions. While the under-reaction to new information friction weakens as local neighborhood size increases, the under-imitation friction appears to intensify when local neighborhoods are large.

Can Under-Imitation Be Rational? While we have established that the two behavioral frictions are distinct, it remains possible that under-imitation is an optimal response to the under-reaction to new information. To evaluate this hypothesis, consider a *Single Aggregator* network with n participants.[46] Assume that (i) n is even ; (ii) every non-aggregator i has at most $\frac{n}{2} - 1$ direct neighbors, i.e., $b(i) < \frac{n}{2}$, and (iii) every two non-aggregators i and j are either not linked, i.e., $ij \notin E$, or they share exactly the same set of neighbors, that is, $B(i)\backslash\{j\} = B(j)\backslash\{i\}$. Note that the Star, the Connected Spokes and the One Gatekeeper networks satisfy these properties. Property (i) introduces ties, property (ii) guarantees non-aggregators never know the majority of private signals for sure already after the first round and property (iii) guarantees that the second round guesses of non-aggregators add no information to their neighbors. Following Findings 6 and 7 and Footnote 42, assume, in addition, that (iv) all subjects guess correctly in the first round, (v) the aggregator, denoted by A, never switches in the second round when her private signal coincides with the majority of first round guesses, (vi) the aggregator does not switch in the second round when her private signal coincides with the minority of first round guesses with probability $\alpha \in (0, 1]$, and (vii) the aggregator does not switch in the second round when there is a tie in the first round guesses with probability $\beta \in [0, 1]$.

Whenever the aggregator switches between round 1 and round 2, their second round guess is surely correct, therefore, in these cases, imitation is optimal. If the aggregator does not switch it might be that her private signal coincides with the majority of first round guesses or there is

[46] We focus here on *Single Aggregator* networks since the Complete network is not expected to exhibit imitation and the *Cluster(s)* networks suffer from structural frictions that may over complicate the discussion (see Section 4.4).

a tie (and then imitation is optimal) or, alternatively, that her private signal coincides with the minority of first round guesses and she decided not to switch (maybe due to under-reaction to new information concerns). When no switch is observed, a Bayesian non-aggregator agent i uses the $b(i)+1$ first round guesses she observed and the fact that the aggregator did not switch, to evaluate the conditional probability that the aggregator's second round guess is incorrect. Claim 1 shows that doubts should emerge only if the aggregator was within agent i's local minority in the first round. The claim's proof is relegated to Section C.1 of the Theoretical Appendix.

Claim 1. *A Bayesian non-aggregator agent i imitates agent A if either (i) the aggregator switched between round 1 and round 2, i.e., $a_A^1 \neq a_A^2$, or (ii) the aggregator did not switch, and their initial guess was not in the first-round minority within agent i's local neighborhood, i.e., $a_A^1 = a_A^2$ and $|j \in B(i) \cup \{i\}|s(j) = s(A)| \geq |j \in B(i) \cup \{i\}|s(j) \neq s(A)|$. If the aggregator did not switch between round 1 and round 2 and their initial guess was in the first-round minority within agent i's local neighborhood, then there exist values of α and β for which imitation is not optimal for agent i.*

Clearly, the aggregator can never be in the local minority of a leaf, so leaf agents should always imitate. Thus, the star network should not exhibit any under-imitation. In Section C.2 of the Theoretical Appendix, we compute the minimal values of α that make imitation suboptimal in the Connected Spokes and One Gatekeeper networks. These values depend on the non-aggregator's position, the size of the local minority, and β. Using the empirical values of α and β,[47] we find that imitation is always optimal in the Connected Spokes network. In the One Gatekeeper network, imitation is optimal when at least two non-aggregators in the clique guessed like the aggregator in round 1. Therefore, under the empirical values of α and β, the only case in which imitation is not optimal for Bayesian non-aggregators in single aggregator networks is when all of the following hold: the agent is a clique member in the One Gatekeeper network, the aggregator does not revise their guess between rounds 1 and 2, and at most one other clique member guessed similarly to the aggregator in round 1. Hence, the theoretical prediction implies extremely low rates of under-imitation—yet observed rates in the laboratory are substantially higher. We conclude that under-imitation cannot be explained as a rational Bayesian response to under-reaction to new information.

Regression Analysis of Imitation Table 7 analyzes the determinants of third-round imitation, incorporating subjects' type, their agreement with the influencer in round 2, the influencer's behavioral change between rounds 1 and 2, and features of the local environment. Regressions (4)-(6) examine imitation behavior separately by position: leafs, cluster members in *Single Aggregator* networks, and cluster members in *Cluster(s)* networks. Regressions (1)-(3) pool all positions to exploit variation in the ratio of the subject's local neighborhood size to that of the influencer—a variable omitted from the position-specific regressions due to limited within-group variation.

[47]In the data, $\alpha = 43.5\%$ for the Connected Spokes network and $\alpha = 45\%$ for the One Gatekeeper network. $\beta = 75\%$ for the Connected Spokes network. Since we observe no single aggregators facing ties after the first round in the One Gatekeeper network, we set here $\beta = 75\%$ as well (also consistent with Footnote 42).

	Dependent Variable					
	Round 3 guess matches round 2 guess of the influencer					
Regression Number	(1)	(2)	(3)	(4)	(5)	(6)
Network Type		All Networks		All Networks	*Single Aggregators*	*Cluster(s)*
Node Types Included	leafs clusters	leafs clusters	leafs clusters	leafs	 clusters	 clusters
Constant	0.928***	0.931***	0.943***	0.882***	0.981***	0.964***
	(0.0186)	(0.0187)	(0.0257)	(0.0327)	(0.0302)	(0.0257)
Individual Controls						
Gender	0.0155	0.0155	0.0188	0.0474**	-0.0133	0.00315
	(0.0115)	(0.0115)	(0.0115)	(0.0207)	(0.0197)	(0.0147)
Probability Matching	-0.0603***	-0.0603***	-0.0548***	-0.0919***	-0.0547**	-0.0193
	(0.0140)	(0.0140)	(0.0140)	(0.0276)	(0.0216)	(0.0196)
Risk Aversion	0.0502**	0.0506**	0.0469**	0.0514	0.0711	0.0212
	(0.0240)	(0.0241)	(0.0238)	(0.0413)	(0.0437)	(0.0303)
Incorrect R1 Guess	-0.105***	-0.105***	-0.106***	-0.129***	-0.0453	-0.133***
	(0.0188)	(0.0188)	(0.0189)	(0.0312)	(0.0314)	(0.0333)
Influencer Round 2 Status						
Disagree with Influencer	-0.591***	-0.588***	-0.511***	-0.478***	-0.730***	-0.824***
	(0.0194)	(0.0223)	(0.0296)	(0.0310)	(0.0371)	(0.0311)
Influencer Switch R1 to R2	-0.000420	-0.00696	-0.0483**	-0.0265	-0.0487**	0.00900
	(0.0104)	(0.0110)	(0.0188)	(0.0208)	(0.0194)	(0.0141)
Disagree with Influencer × Influencer Switch	0.0819*** (0.0308)	0.0675** (0.0326)	0.160*** (0.0458)	0.111** (0.0493)	0.166*** (0.0558)	-0.0840* (0.0473)
Minority Status						
In R2 Minority		-0.0402*	-0.0440*		-0.0764***	-0.0663
		(0.0237)	(0.0234)		(0.0286)	(0.0406)
In R2 Minority × Disagree with Influencer		0.0196 (0.0362)	0.106*** (0.0361)		0.141*** (0.0497)	0.271*** (0.0636)
In R2 Minority × Influencer Switch		0.0718* (0.0411)	0.122*** (0.0398)		0.117** (0.0522)	0.0639 (0.0772)
Network Features						
Ratio			-0.0465			
			(0.0491)			
Ratio × Influencer Switch			0.0731**			
			(0.0289)			
Ratio × Disagree with Influencer			-0.317*** (0.0519)			
Ratio × Influencer Switch × Disagree with Influencer			-0.314*** (0.0902)			
R-squared	0.432	0.433	0.451	0.322	0.500	0.623
# of Observations	4,521	4,521	4,521	1,933	1,292	1,296
# of Clusters	721	721	721	360	244	237
# of Session FEs	36	36	36	18	12	12

Table 7: Determinants of third-round imitation

Notes: All regressions are linear, with standard errors clustered at the participant level and session fixed effects included. Regs (1)-(3) use a pooled sample of all non-aggregators in the *Single Aggregator* networks, leafs in the Symmetric Core–Periphery network, and non-connectors in the Two Cores networks. Reg (4) includes leafs in the Star, One Gatekeeper, and Symmetric Core–Periphery networks. Reg (5) includes cluster members in the Connected Spokes and One Gatekeeper networks. Reg (6) includes non-connectors in the Two Cores networks. *Disagree with influencer* is an indicator for whether the subject's round 2 guess differs from their influencer's round 2 guess. *Influencer switch* indicates whether the influencer changed their guess between rounds 1 and 2. *In R2 minority* indicates whether the subject's round 2 guess was not the local majority in their neighborhood. *Ratio* is defined as the number of the subject's direct neighbors divided by the number of the influencer's direct neighbors. Individual controls include the risk attitude measure, the probability matching indicator, the indicator of sub-optimal first round guess, and gender. Robustness checks are reported in Section D.3 of the Empirical Appendix.

The regressions in Table 7 yield several notable findings. Throughout the analysis we focus on subjects who, under the Bayesian model, are expected to imitate their influential neighbor—specifically, those who are not probability matchers and that guessed correctly in round 1. First, by regression (1), when these subjects agree with the influencer in round 2, they maintain their guess in 92.8% of cases, consistent with imitation. However, when imitation requires switching—i.e., when their round 2 guess differs from the influencer's—imitation drops sharply to 33.7% (assuming the influencer submitted the same guess in round 1 and round 2). We identify three main factors that shape the extent of this drop: (i) the behavior of the local neighborhood, (ii) the behavior of the influencer, and (iii) structural features of the influencer's network position.

Regressions (5) and (6) reveal how the behavior of the local neighborhood affects imitation. Consider the case where the influencer does not switch between rounds 1 and 2. When non-influencers agree with the influencer, and thus do not need to switch under the Bayesian model, imitation rates are well over 96% when the subject is in the local majority in round 2. These rates drop slightly to about 90% when in the local minority. The role of the local environment becomes more pronounced when the subject disagrees with the influencer. In this case, imitation is rare when the subject is in the local majority—just 25% in *Single Aggregator* networks and 14% in *Cluster(s)* networks. However, when the subject is in the local minority—i.e., most of their neighbors agree with the influencer—imitation improves. This improvement accounts for 6.5 percentage points of under-imitation in *Single Aggregator* networks and at least 20.5 percentage points in *Cluster(s)* networks. Put differently, when subjects are in the local minority and agree with the influencer, they rarely switch to match the local majority—imitation rates drop by 7.6 percentage points in *Single Aggregator* networks and insignificantly in *Cluster(s)* networks. But when they disagree with the influencer, being in the local minority increases switching rates, especially among non-connectors in the *Cluster(s)* networks. Notably, the high rates of under-imitation among those who disagree with the influencer are inconsistent with the Bayesian model, while the frequent refusal to conform to the local majority among those who agree with the influencer stands in sharp contrast to the naïve model.[48]

Both the Bayesian and naïve models predict that once the influencer's second-round guess is known, their first-round guess should be irrelevant for determining the subject's third-round decision. However, regression (4) shows that leaf subjects who disagree with the influencer are 11.1 percentage points more likely to imitate when they observe that the influencer switched between rounds 1 and 2—accounting for 23.2% of the under-imitation effect. Regression (5) reveals a similar pattern among cluster members in *Single Aggregator* networks: when they disagree with the influencer, observing a switch increases imitation rates by 11.7 percentage points when they are in the local majority and by 23.4 percentage points when they are in the local minority.

To assess the effect of the influencer's position on imitation, regression (3) includes the variable *Ratio*, defined as the size of the subject's local neighborhood divided by that of the influencer (i.e.,

[48]Consider third-round decisions by subjects in non-leaf positions who followed both models in the first two rounds and for whom both models yield clear third-round predictions (462 observations). A direct, uncontrolled comparison shows that when both models predict no switch, only 1.8% of subjects switch. When the naïve model predicts a switch but the Bayesian model does not, 5.1% switch. In contrast, when the Bayesian model predicts a switch and the naïve model does not, 32.8% switch. Even when both models predict switching, only 42.1% of subjects switch.

the subject's degree centrality divided by the influencer's degree centrality). This measure ranges from $\frac{1}{17}$ for leafs in *Single Aggregator* networks to $\frac{8}{9}$ for non-connectors in the Two Cores with One Link network. The results show that when subjects disagree with the influencer, imitation rates decline as the influencer's informational advantage diminishes—that is, as *Ratio* increases. In addition, regression (3) indicates that observing the influencer switch promotes imitation among subjects who are much less connected than the influencer (e.g., non-aggregators in *Single Aggregator* networks), but reduces imitation among subjects with similarly sized local neighborhoods (e.g., non-connectors in the Two Cores with One Link network).

In conclusion, third-round behavior of potential imitators deviates from the predictions of both the myopic Bayesian and naïve models. The Bayesian model is undermined by the low imitation rates observed when the subject and influencer disagree in round 2, while the naïve model cannot account for the low frequency of switching to match the round 2 local majority. We argue that under-imitation reflects an excessive over-reaction to concerns that the influencer may have under-reacted to new information. Disagreement in round 2 triggers a reassessment of the influencer's trustworthiness. The observed patterns suggest that imitators rely on two key cues to form this judgment. First, agreement between the influencer and the subject's local majority serves as a credibility cue, increasing imitation. Second, when the influencer switches between rounds 1 and 2, it signals responsiveness to new information, which also encourages imitation. The effect of this latter signal depends on the influencer's informational advantage: the larger the gap, the stronger the response. Together, these two cues raise imitation rates by 30 percentage points in *Single Aggregator* networks and nearly 20 percentage points in *Cluster(s)* networks.

Finding 8. *Participants tend to imitate neighbors with superior information, but do so far less frequently than optimal when their second-round guess differs from that of the influencer. This pattern of under-imitation cannot be explained by rational concerns about under-reaction to new information or by naïve behavior. Instead, we argue that it reflects an irrational and excessive over-reaction to mis-aggregation concerns. Two cues appear to improve imitation rates: agreement between the influencer and the subject's local majority and switching by the influencer between rounds 1 and 2 when the influencer holds a clear informational advantage.*

5.4 Guesses Beyond the Third Round

Finding 5 highlights that behavior in the first three rounds largely determines network outcomes, with ALI rates stabilizing from round 4 onward. In practice, 41% of subjects never switched after round 4, and 84% switched in at most two games. The positional analysis in Panel A of Table 20 (Section D.4 of the Empirical Appendix) reveals that, for most positions, no switches occur after round 3 in at least 80% of cases—the exceptions being positions affected by structural frictions.

5.5 Final Round Guesses

Table 8 reports regression results on the determinants of correct final-round guesses, analyzed by network position. Recall that the final round was determined endogenously and was not distinctively incentivized.

Across all positions, early-round mistakes emerge as a consistent and powerful predictor of incorrect final guesses: misreporting the private signal in round 1 (for non-influencers in incomplete networks), mis-aggregating local information in round 2 (for non-leaf positions), or failing to imitate optimally in round 3 (for potential imitators). These long-lasting negative effects highlight the central role of behavioral frictions—specifically, under-reaction to new information and under-imitation—in shaping individual decisions.

As discussed in Section 5.4, late-round switching was relatively rare. Nevertheless, for participants who made early mistakes, late switches partially mitigated the damage—recovering between 55% and 80% of the initial loss. By contrast, for those who made no early error, late switches tended to reduce the likelihood of a correct final guess.

Regressions (1) and (2) offer an illuminating comparison between the participants in the Complete network and the aggregators in the *Single Aggregator* networks, echoing patterns seen in Figure 2. In Regression (1), a larger local minority size in round 1 significantly reduces the probability of a correct final guess in the Complete network. This effect is absent in Regression (2), despite both samples being limited to participants who observe the full network. Given that most participants act on their private signals in round 1 (see Finding 6), and that both the Complete network participants and the aggregators in *Single Aggregator* networks observe all others, a larger *R1 Local Minority Size* implies lower initial signal quality. Thus, while final guesses in the Complete network are sensitive to the quality of initial signals,[49] the final performance of aggregators in *Single Aggregator* networks appears unaffected. This discrepancy is not accounted for by either the Bayesian or the naïve models and suggests the influence of an unobserved factor—possibly related to the connectivity of other agents—on the aggregation process.

Regression (3) reveals another noteworthy result. In Section 3.4, we defined a surmountable structural friction and applied it to the Two Cores networks: when connectors from opposite cliques disagree in round 2, myopic Bayesian agents should, in some cases, optimally switch only in later rounds. Regression (3) provides two relevant observations. First, when signal quality is poor—as indicated by a large *R1 Local Minority Size*—connectors are significantly less likely to guess correctly in the final round when they disagree. Second, in cases of round 2 disagreement, late switches by the connectors fail to improve their accuracy. This suggests that such switching behavior is not driven by sophisticated Bayesian reasoning.

Finally, while most networks in the experiment feature clusters of similar size, limiting our ability to study the role of local environment size systematically, the Connected Spokes networks offer a useful exception. They include both small clusters (three non-aggregators plus one aggregator) and large clusters (four non-aggregators plus one aggregator). Regression (5) shows that non-aggregators

[49] Regression (5) exhibits similar sensitivity by cluster members that are non-influencers.

	Reg (1) Complete	Reg (2) Aggregators	Reg (3) Connectors	Reg (4A) Leafs Only Core Periphery	Reg (4B) Leafs Only All Others	Reg (5) Clusters with Influencers
Constant	1.082***	1.008***	1.142***	0.912***	0.826***	0.998***
	(0.0399)	(0.0854)	(0.0807)	(0.0557)	(0.0320)	(0.0257)
Individual Controls						
Probability Matching	-0.00365	0.00142	-0.0885**	0.0335	0.0244	0.0200
	(0.0199)	(0.0542)	(0.0431)	(0.0588)	(0.0307)	(0.0165)
Gender	-0.0169	0.0246	-0.0337	-0.0273	0.0237	-0.0153
	(0.0142)	(0.0324)	(0.0349)	(0.0391)	(0.0182)	(0.0127)
Risk Aversion	0.0247	-0.142*	0.0162	0.0443	0.00465	0.0369
	(0.0265)	(0.0827)	(0.0682)	(0.0642)	(0.0403)	(0.0294)
Initial Behavior						
Wrong R1 Guess	-0.0160	-0.111	-0.0422	-0.173***	-0.0952**	-0.0847***
	(0.0402)	(0.0879)	(0.0575)	(0.0646)	(0.0407)	(0.0315)
Wrong R2 Guess	-0.931***	-0.917***	-0.573***	-0.0335	-0.0215	-0.312***
	(0.0284)	(0.0607)	(0.0967)	(0.145)	(0.0898)	(0.0260)
Wrong R3 Guess				-0.399***	-0.527***	-0.331***
				(0.0780)	(0.0384)	(0.0288)
Late Switching Behavior						
Switched in R3+	-0.0958*	-0.294***	0.0254			
	(0.0518)	(0.0976)	(0.0752)			
Wrong R2 Guess	0.816***	0.862***	0.390**			
× Switched in R3+	(0.0896)	(0.182)	(0.154)			
Switched in R4+				-0.126**	-0.0887***	-0.110***
				(0.0638)	(0.0337)	(0.0301)
Wrong R3 Guess				0.354***	0.475***	0.294***
× Switched in R4+				(0.132)	(0.0633)	(0.0509)
Local Network Information						
R1 Local Minority Size	-0.340***	0.118	-0.0962			-0.218***
	(0.0918)	(0.182)	(0.135)			(0.0483)
Core Connectors Disagree			0.00992			
			(0.0756)			
Core Connectors Disagree			-0.699**			
× R1 Local Minority Size			(0.278)			
Core Connectors Disagree			0.00933			
× Switched in R3+			(0.0943)			
Influencer Switched in R3+				-0.105***	-0.0159	-0.0216
				(0.0376)	(0.0313)	(0.0183)
Network Structure						
Three-Connecting Node			-0.0833**			
			(0.0405)			
Connected Spoke Small Cluster						-0.0457*
						(0.0242)
R-squared	0.578	0.722	0.353	0.107	0.182	0.255
# of Observations	684	159	318	522	1,411	2,900
# of Clusters	106	128	165	119	241	484
# of Session FEs	5	18	12	6	12	24

Table 8: Determinants of Last Correct Guess

Notes: All regressions are linear, with standard errors clustered at the participant level and session-game fixed effects included. Reg (1) uses data from the Complete network; (2) from aggregators in *Single Aggregator* networks; (3) from connectors in Two Cores networks; (4A) from leafs in the Symmetric Core–Periphery network; (4B) from leafs in the Star and One Gatekeeper networks; and (5) from non-connectors in Two Cores networks, non-aggregator cluster members in the One Gatekeeper network, and non-aggregators in the Connected Spokes network. The dependent variable, *Last Correct Guess*, equals 1 if the participant guessed correctly in the final round. *Wrong Rx Guess* equals 1 if the participant guessed not according to the myopic Bayesian model in round x. *Switched in Ry+* equals 1 if the participant switched at any round $t \geq y$ relative to round $y-1$. *R1 Local Minority Size* is the fraction of minority guesses in the participant's local neighborhood in round 1. *Core Connectors Disagree* equals 1 whenever there is no unanimity amongst the connectors in round 2 in the Two Cores networks. *Influencer Switched in R3+* equals 1 if the influencer switched at any round $t \geq 3$ compared to round 2. *Three-Connecting Node* indicates whether the participant is one of the three connectors in the Two Cores with Three Links network. *Connected Spoke Small Cluster* indicates assignment to a small cluster in the Connected Spokes network. Individual controls include risk attitude, probability matching, and gender. Robustness checks appear in Section D.5 of the Empirical Appendix. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

in the smaller clusters are 4.57 percentage points less likely to guess correctly in the final round. We interpret this as anecdotal evidence supporting the view that larger local environments facilitate more accurate final guesses.

Finding 9. *Subjects rarely revise their guesses after the third round, making early under-reaction to new information and under-imitation persistent frictions with lasting effects on performance. In the Complete network, final guesses are sensitive to the quality of initial signals, whereas aggregators in Single Aggregator networks appear unaffected. In the Two Cores networks, we find no evidence of sophisticated Bayesian reasoning.*

6 Intervention: Mitigating the Behavioral Frictions

Our position-level analysis reveals that two behavioral frictions—under-reaction to new information and under-imitation—significantly hinder participants' ability to correctly identify the state of the world. This section presents a follow-up experiment showing that reducing the amount of information available to specific participants can partially mitigate both under-reaction and under-imitation.

6.1 Design

Learning in *Single Aggregator* networks relies on the aggregator's ability to accurately aggregate first-round signals and relay the result to others. Sections 4 and 5 show that these networks are particularly prone to under-reaction to new information. To mitigate this friction, we implement a simple intervention: withholding the aggregator's private signal to reduce the risk of early mis-aggregation. All other participants receive partially informative signals and are explicitly informed that the aggregator receives none. We assess the intervention's effectiveness using the ALI metric and position-level accuracy.

We implement this intervention in the One Gatekeeper network, which features positional heterogeneity among non-aggregators. Specifically, we conduct six additional experimental sessions that replicate the original six One Gatekeeper sessions.[50] In each game of the new treatment, non-aggregator participants received the same private signal as their counterparts in the original sessions, while the aggregator received the message: "In Round 1 you received NO SIGNAL." All participants were explicitly informed that the aggregator received this message while they themselves received a private informative signal.[51] We refer to these new sessions as *One Gatekeeper Scripted*. By holding initial signals fixed for all except the aggregator, any observed behavioral differences between the two treatments can be attributed to the aggregator's lack of private information.

[50] Due to the COVID-19 pandemic, these sessions were conducted online rather than in a physical lab. The subject pool consisted of 120 undergraduate students at The Ohio State University. Experimental instructions are provided in Section A.5 of the Empirical Appendix.

[51] In a similar design, Choi et al. (2005, 2012) use three-person networks in which agents receive a private signal with probability $q < 1$. However, in their setup, participants do not know whether others are informed or not.

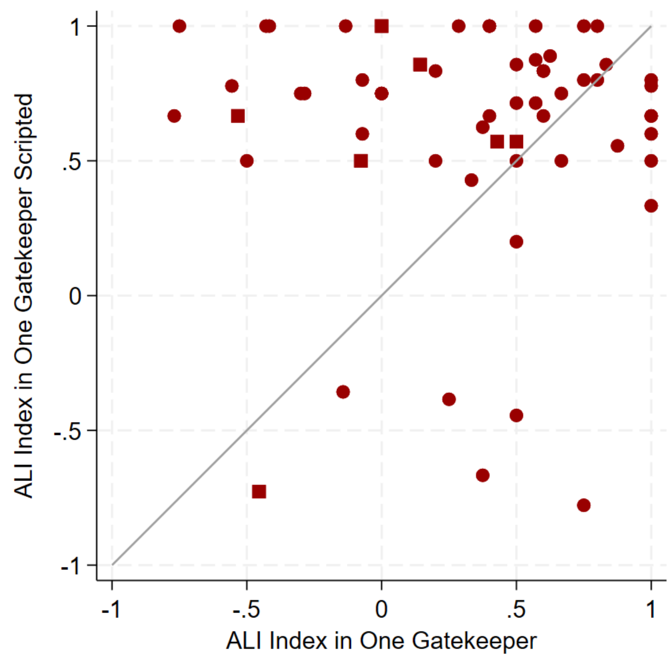


Figure 6: One Gatekeeper vs. One Gatekeeper Scripted

Notes: Each dot represents a matched pair of games. The horizontal axis reports the ALI value from the game played under a standard One Gatekeeper session, while the vertical axis reports the ALI from the corresponding game in a Scripted session. Circles indicate pairs where both games ended in fewer than 50 rounds (i.e. converged); squares indicate pairs where the original game ended in more than 50 rounds (i.e. did not converge). Two matched pairs of games are excluded from the figure because the original game featured a perfectly balanced signal distribution (nine signals of each state).

6.2 Analysis

Figure 6 plots matched game pairs, with ALI from the original One Gatekeeper sessions on the x-axis and from the corresponding Scripted sessions on the y-axis. Dots above the 45-degree line indicate more effective information aggregation in the Scripted sessions; dots below indicate the opposite. The figure shows that withholding a private signal from the aggregator—while holding all else constant—improved learning in the One Gatekeeper network.[52]

Two factors explain this improvement. First, aggregators without a private signal performed better in the second round: correct guesses rose from 86% in original sessions to 90% in Scripted ones. This gain came primarily when the aggregator's initial guess was in the minority. In such cases, switching to the correct answer in round 2 increased from 59% to 90% ($p = 0.015$). By contrast, when the first-round guess aligned with the majority, accuracy remained high and similar across treatments (100% vs. 90%, $p = 0.056$). Thus, over 70% of aggregation errors in the original sessions—among aggregators initially in the minority—were eliminated when the private signal was

[52] A binomial probability test rejects the null hypothesis that dots are equally likely to fall above or below the 45-degree line ($p \approx 0.001$). Figure 6 in Section C.4 of the Empirical Appendix replicates this using the ILI metric ($p \approx 0.002$).

Panel A: Regression Estimates

	Reg (1) All Non-Aggregators	Reg (2) Leafs Only	Reg (3) Non-Aggregator Cluster Roles: All	Reg (4) Non-Aggregator Cluster Roles: R2 Majority	Reg (5) Non-Aggregator Cluster Roles: R2 Minority
Constant	0.957***	0.967***	0.956***	0.971***	0.826***
	(0.0295)	(0.0441)	(0.0368)	(0.0280)	(0.136)
Individual Controls					
Gender	0.0150	0.0208	0.00139	-0.0140	0.0889
	(0.0182)	(0.0257)	(0.0224)	(0.0188)	(0.0735)
Probability Matching	-0.0492**	-0.0718**	-0.0259	-0.00615	-0.110
	(0.0213)	(0.0306)	(0.0258)	(0.0155)	(0.0777)
Risk Aversion	-0.0421	-0.0406	-0.0483	0.0223	-0.360**
	(0.0364)	(0.0520)	(0.0424)	(0.0344)	(0.161)
Incorrect R1 Guess	-0.124***	-0.170***	-0.0680	-0.000118	-0.142*
	(0.0349)	(0.0582)	(0.0482)	(0.0446)	(0.0828)
Aggregator Information					
Disagree with Aggregator	-0.594***	-0.531***	-0.672***	-0.687***	-0.536***
	(0.0422)	(0.0576)	(0.0456)	(0.0969)	(0.101)
Aggregator Switch R1 to R2	0.0237	-0.0205	0.0806**	0.0414*	0.112
	(0.0319)	(0.0448)	(0.0385)	(0.0211)	(0.115)
Disagree with Aggregator	0.143***	0.101*	0.181**	0.0688	0.213*
× Aggregator Switch	(0.0452)	(0.0582)	(0.0751)	(0.132)	(0.112)
Scripted Treatment					
Scripted Flag	0.0283	-0.00184	0.0655**	0.0236	0.260*
	(0.0214)	(0.0300)	(0.0284)	(0.0273)	(0.135)
Scripted Flag	0.115**	0.149**	0.0356	-0.0935	0.0781
× Disagree with Aggregator	(0.0560)	(0.0714)	(0.0775)	(0.148)	(0.130)
Scripted Flag	-0.0636	-0.0279	-0.115**	-0.0476	-0.215
× Aggregator Switch	(0.0446)	(0.0622)	(0.0519)	(0.0308)	(0.151)
Panel B: Scripted and Disagreement Contrast					
Scripted Flag + Scripted Flag × Disagree with Aggregator	0.144**	0.147*	0.101	-0.070	0.338***
	(0.062)	(0.079)	(0.076)	(0.143)	(0.106)
Observations	1,887	999	888	688	200
# of Matches	111	111	111	105	83
# of Participants	242	239	242	234	134

Table 9: Imitation in the Third Round: One Gatekeeper vs. One Gatekeeper Scripted

Notes: All regressions in Panel A are linear, with standard errors clustered at the participant level and no fixed effects included. The sample includes 51 standard One Gatekeeper games that converged and were not tied, and all 60 One Gatekeeper Scripted games. The dependent variable, *Correct Third Round Guess*, equals 1 if the participant's third-round guess matched the aggregator's second-round guess. *Disagree with Aggregator* equals 1 if the participant's second-round guess differed from the aggregator's second-round guess. *Aggregator Switch R1 to R2* equals 1 if the aggregator changed their guess between rounds 1 and 2. *Scripted Flag* equals 1 for games played in a Scripted session. Individual controls include the risk attitude measure, the probability matching indicator, the indicator of sub-optimal first round guess, and gender. Panel B uses the results exhibited in Panel A to calculate the difference between rates of imitation for participants in the scripted session who disagree with the aggregator in round 2 and participants in the unscripted session who disagree with the aggregator in round 2. Robustness checks appear in Section E.1 of the Empirical Appendix. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

withheld. This suggests that removing the private signal from the aggregator directly mitigated under-reaction to new information.

Second, withholding the aggregator's private signal also changes the behavior of others in the network by increasing imitation—an indirect effect of the intervention. Recall that in the One

	Single Aggregators	Leafs Only	Clusters Only
Constant	0.694***	0.900***	0.752***
	(0.220)	(0.0415)	(0.0842)
Individual Controls			
Gender	-0.0474	0.0236	0.0211
	(0.0596)	(0.0265)	(0.0270)
Probability Matching	-0.0268	-0.0893**	-0.0160
	(0.0631)	(0.0357)	(0.0342)
Risk Aversion	0.0172	-0.0522	-0.0236
	(0.0891)	(0.0488)	(0.0393)
Incorrect R1 Guess	-0.251	-0.196***	-0.224***
	(0.314)	(0.0755)	(0.0736)
Switched in R3+	-0.189	-0.00105	0.0105
	(0.140)	(0.0357)	(0.0408)
Information and R1 Minority			
Signal Wrong	-0.346***	-0.372***	-0.352***
	(0.128)	(0.0607)	(0.0598)
Size of R1 Majority	0.522		0.203**
	(0.332)		(0.0942)
Aggregator Signal			-0.000320
Matches R1 Majority			(0.0445)
Scripted Treatment			
Scripted Flag	-0.0695	0.00446	0.0232
	(0.0531)	(0.0274)	(0.0483)
Scripted Flag	0.329**	0.176**	0.166**
× Signal Wrong	(0.153)	(0.0736)	(0.0700)
Scripted Flag			0.0631
× Agg Signal Match R1 Major			(0.0507)
R-squared	0.181	0.138	0.182
Observations	111	999	888
Clusters	88	239	242

Table 10: Final Guess Accuracy: One Gatekeeper vs. One Gatekeeper Scripted

Notes: All regressions are linear, with standard errors clustered at the participant level. The sample includes 51 standard One Gatekeeper games that converged and were not tied, and all 60 One Gatekeeper Scripted games. The dependent variable, *Correct Final Guess*, equals 1 if the participant's final-round guess was accurate. *Signal Wrong* equals 1 if the participant's private signal was incorrect. For the aggregator in scripted games we use the signal in the corresponding unscripted game. *Size of R1 Majority* is the fraction of majority guesses in the participant's local neighborhood in round 1. *Aggregator Signal Matches R1 Majority* equals 1 if the aggregator's signal matched the local majority in the first round. For scripted games we use the aggregator's signal in the corresponding unscripted game. *Scripted Flag* equals 1 for games played in a Scripted session. Individual controls include the risk attitude measure, the probability matching indicator, the indicator of sub-optimal first round guess, the indicator of late switching and gender. Robustness checks appear in Section E.2 of the Empirical Appendix. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Gatekeeper network, it is optimal for all non-aggregators to imitate the aggregator from round three onward.[53] Table 9 compares third-round imitation across original and Scripted sessions. In Regression (2), original sessions show that when a leaf's second-round guess disagrees with the aggregator's, imitation drops by 53.1 percentage points.[54] Panel B shows this drop is significantly smaller in Scripted sessions—under-imitation reduces by over 25%. A similar pattern holds for

[53]Result 2 in Section B of the Theoretical Appendix extends to the case where the aggregator lacks a private signal, provided the aggregator is unique.

[54]This matches the effect in Regression (4) of Table 7. Effects of observing the aggregator switch between rounds are also consistent.

cluster members in the local minority after round 2 that disagree with the aggregator: as Regression (5) shows, their imitation drop of 53.6 percentage points in original sessions is reduced by over 60% in Scripted ones. By contrast, Regression (4) shows no effect among cluster members in the local majority who disagree with the aggregator—imitation remains low in both original and scripted sessions.[55]

Table 10 analyzes the long-run effects of the intervention, revealing two key findings. First, for aggregators, receiving no private signal is as effective as receiving a correct one—and significantly better than receiving a wrong one. This highlights a striking result: withholding information from a fully connected agent improves her long-term performance. Second, for non-aggregators, the negative impact of receiving an incorrect signal is nearly halved when the aggregator lacks a signal. As shown earlier, this is driven by greater aggregator accuracy and increased imitation, especially by leafs and cluster minorities.

The reduced under-imitation friction observed in the One Gatekeeper Scripted sessions, along with third-round imitation patterns observed in original One Gatekeeper sessions (Table 7), points to trust as a key driver of imitation. As noted in Finding 7, imitators may over-doubt the aggregator's judgment due to common under-reaction to new information. However, imitation becomes more likely—when it involves switching from one's own guess—if cues boost confidence in the aggregator. Three cues stand out from our analysis: (1) alignment between the aggregator and the local majority, (2) the aggregator revising their guess between rounds 1 and 2, signaling responsiveness, and (3) knowing the aggregator lacks a private signal, suggesting her guess reflects true information aggregation rather than bias. Each cue can enhance trust in the aggregator and thereby increase the likelihood of imitation.

Finding 10. *Depriving the aggregator from having a signal mitigates under reaction to new information, which in turn increases trust and improves imitation—especially among leafs and second-round minority cluster members. Overall, this intervention significantly enhances information aggregation in One Gatekeeper networks.*

7 Closing the Loop: Back to Network Level Performance

Section 4 shows that aggregation is imperfect even in the Complete network, and its success depends on both network structure and signal quality. *Cluster(s) networks* match Complete network performance when signals are strong but often fail when signals are weak. In contrast, *Single Aggregator networks* perform well with poor initial information but show little improvement as signal quality rises, frequently failing even when most initial signals are correct.

In this section, we explain these network-level patterns using the structural frictions from Section 3.4 and the behavioral frictions from Sections 5.2 and 5.3. In particular, we use insights from the positional-level analysis in Section 5 and the intervention in Section 6 to account for network-level

[55]One might worry that withholding the signal increased the aggregator's salience. But the position- and history-dependent response patterns are inconsistent with a general salience effect.

performance through individual-level behavior. We conclude with a welfare analysis identifying positions with superior average and long-run performance.

7.1 The Complete Network

Result 1 shows that both the Bayesian and naïve models predict absolute learning in the Complete network: every subject should converge to the correct state, yielding ALI = 1. Empirically, Figure 2 confirms that information aggregation is more successful in the Complete network than in any other structure. Yet, performance falls short of the theoretical benchmark. As shown in Panel A of Figure 3, only 16% of games reach full convergence (ALI = 1), and the average ALI across all games is approximately 0.6. Moreover, Figure 2 and Panel B in Figure 3 reveal a strong positive relationship between ALI and the overall quality of private signals.

We attribute these deviations to the behavioral friction of *Under-Reaction to New Information*. By the start of the second round, some subjects recognize that their private signal conflicts with the majority's. Yet in nearly half the cases, they fail to revise their guess (Table 4). The last two regressions in Table 5 show that when a subject is alone in the minority after round 1, the probability of not switching is about 18.5%, rising by roughly 4.8 percentage points for each additional incorrect signal observed. This minority-size sensitivity helps explain why ALI is positively associated with signal quality.

The friction persists beyond round 2. Finding 9 and Regression (1) in Table 8 show that even in the Complete network—where errors are easily detectable—subjects rarely correct mistakes in later rounds, and their final guesses remain sensitive to the size of the round 1 minority. We conclude that performance in the Complete network is constrained by *Under-Reaction to New Information*, which intensifies with weaker private signals and persists across rounds.

7.2 The Single Aggregator Networks

The Bayesian model predicts absolute aggregation (ALI = 1) in *Single Aggregator* networks: the aggregator should correctly combine all signals by round 2, and all non-aggregators should imitate her from round 3 onward (Result 2 in Section B of the Theoretical Appendix and the discussion in Section 3.4). By contrast, the naïve model treats the aggregator as an equally informed peer. As a subject's neighborhood grows, they place less weight on the aggregator's guess, rely excessively on local private signals, and ultimately cause information aggregation failures (Results 3, 4, and 5 in Section B of the Theoretical Appendix).

Empirically, *Single Aggregator* networks perform surprisingly poorly. Panel A of Figure 3 shows that about a quarter of their games exhibit relative information aggregation failures (ALI $<$ 0), and in over half of these, the final majority guess is incorrect. Figure 2 shows that with low-quality signals, the performance of *Single Aggregator* networks matches the Complete network and exceeds *Cluster(s)* networks. However, unlike these structures, performance does not improve with signal quality (Panel B in Figure 3); when signals are mostly correct, *Single Aggregator* networks fall behind both.

As in the Complete network, subjects in *Single Aggregator* networks under-react to new information. Table 4 shows that second-round switching occurs in only 42–61% of cases among aggregators and 31–53% among participants who are neither leafs nor aggregators. When aggregators receive no private signal, however, this friction is largely mitigated: second-round switching failures drop to about 10%. Indeed, Table 10 shows that for aggregators, receiving no signal is as effective as receiving a correct one.

Table 6 shows that non-aggregators imitate infrequently when it requires changing their guess: leaf agents do so in only 45% of such cases, and others in about one-third. Table 8 confirms that these third-round imitation failures have a lasting negative impact on belief accuracy. Claim 1 and the subsequent discussion show that such low imitation rates are not rational responses to the aggregator's under-reaction to new information, while Footnote 48 indicates that they are also inconsistent with the naïve model. The evidence indicates that under-imitation is an excessive and irrational reaction to the aggregator's under-reaction. Tables 7 and 9 show that leaf nodes and second-round local minority agents who disagree with the aggregator imitate more when the aggregator's under-reaction seems unlikely, but this trust-enhancing effect does not extend to second-round local majority agents who disagree with the aggregator.

We conclude that performance in *Single Aggregator* networks is constrained by two behavioral frictions: the aggregator's under-reaction to new information and the non-aggregators' under-imitation. Unlike in the Complete network, the size of the first-round minority does not affect the aggregator's long-run performance (Section 5.5), helping to explain why performance in *Single Aggregator* networks fails to improve with signal quality. One possible explanation is that aggregators in these networks lack the implicit monitoring pressures present in the Complete network.[56]

7.3 The Cluster(s) Networks

Both the Bayesian and naïve models predict that, in most cases, core members in the Symmetric Core–Periphery network aggregate only their own signals, though they differ in their predictions for leafs (Result 6 in Section B of the Theoretical Appendix). Aggregating solely within the core creates a *structural friction* when the core majority does not align with the global majority. In Two Cores with One Link and Two Cores with Three Links networks, frictions arise when the internal majorities of the two cores are misaligned. Under the Bayesian model, these are *surmountable structural frictions*: highly sophisticated behavior by the connectors can always lead to correct information aggregation (Results 7 and 9 in Section B of the Theoretical Appendix).

Experimentally, Figure 2 shows that when the aggregate quality of private signals is low, ALI is

[56] Social facilitation theory—particularly the concept of evaluation apprehension—suggests that concern about being judged by others can influence behavior and performance (Cottrell (1972), Aiello and Douthitt (2001) and Guerin (2010)). We are not aware of applications of this theory in experimental social networks, but in our setting, more connected neighbors may be perceived as more judgmental and knowledgeable. Since neighbor connectivity should not affect second-round guesses, we test the monitoring hypothesis by adding the maximum degree centrality among a subject's neighbors to a regression where the dependent variable is correct round 2 guess (final column of Table 5). The coefficient (0.186) is positive but statistically insignificant ($t = 1.26$), likely due to limited variation—only five values—and absorption by session fixed effects. Therefore, we omit it from the reported regression.

substantially lower in *Cluster(s)* networks than in Complete or *Single Aggregator* networks. However, as signal quality improves, ALI rises, eventually matching the performance of the Complete network.

Figure 4 illustrates the role of both structural and surmountable structural frictions in *Cluster(s)* networks. Panel A highlights the impact of signal quality on the Symmetric Core–Periphery network, compared with the One Gatekeeper network: ALI improves as core signal quality increases. All cases of substantial incorrect learning ($ALI < -0.25$) feature a slim majority in the core, where peripheral information struggles to enter due to potential ties. Panel B shows that misaligned core majorities reduce ALI in Two Cores networks, particularly in the Two Cores with One Link network. Both types of structural frictions are far more likely when the overall signal distribution is weak.[57]

In addition, as in other networks, subjects in *Cluster(s)* networks under-react to new information. Table 4 shows that non-connectors switch in 52%–58% of cases—similar to rates in the Complete network—while connectors switch more often, at 65%–71%. These higher rates of switching among connectors support our speculative hypothesis that implicit monitoring pressures improve second-round aggregation.

Cluster(s) networks demonstrate that *Under-imitation* depends on the network structure. Table 6 shows that while leafs in the Symmetric Core–Periphery network imitate in 60% of cases, imitation rates in the Two Cores networks are substantially lower—only 21%–25%. Regression (3) in Table 7 highlights that when subjects disagree with the influencer, imitation rates fall as the influencer's informational advantage diminishes. This reluctance to imitate among non-connectors reflects the substantial overlap between their neighborhood and that of the connector, whose only additional links are to counterparts in the other core.

We conclude that performance in *Cluster(s)* networks is hindered by both structural and behavioral frictions. Structurally, segregated groups struggle to integrate outside information. Behaviorally, imitation rates are sensitive to the informational advantage of the influencer. These frictions are particularly damaging when signal quality is weak, leading to poor performance. However, they are largely neutralized when signal quality is high. As a result, since under-reaction to new information in *Cluster(s)* networks is comparable in magnitude to that observed in the Complete network when private signals are of high quality, the performance of *Cluster(s)* networks approaches that of the Complete network.

7.4 Performance by Network Positions: Winners and Losers

In Sections 7.1 - 7.3, we used structural and behavioral frictions to interpret the empirical patterns in the aggregate performance of networks. In this section, we use these frictions to account for positional heterogeneity in performance.

Figure 7 shows the frequency of correct guesses, calculated over all rounds (overall performance) and for the final guess (long-term performance). These measures capture welfare differences

[57] For example, we define strong-signal games as those in which no more than four agents receive incorrect signals (see Section 4.1). With strong signals, core majorities are necessarily aligned in the Two Cores networks, and a clear majority in the core of the Symmetric Core–Periphery network is guaranteed unless all incorrect signals are assigned to core members. Thus, the likelihood of structural frictions in *Cluster(s)* networks is negligible when signals are strong.

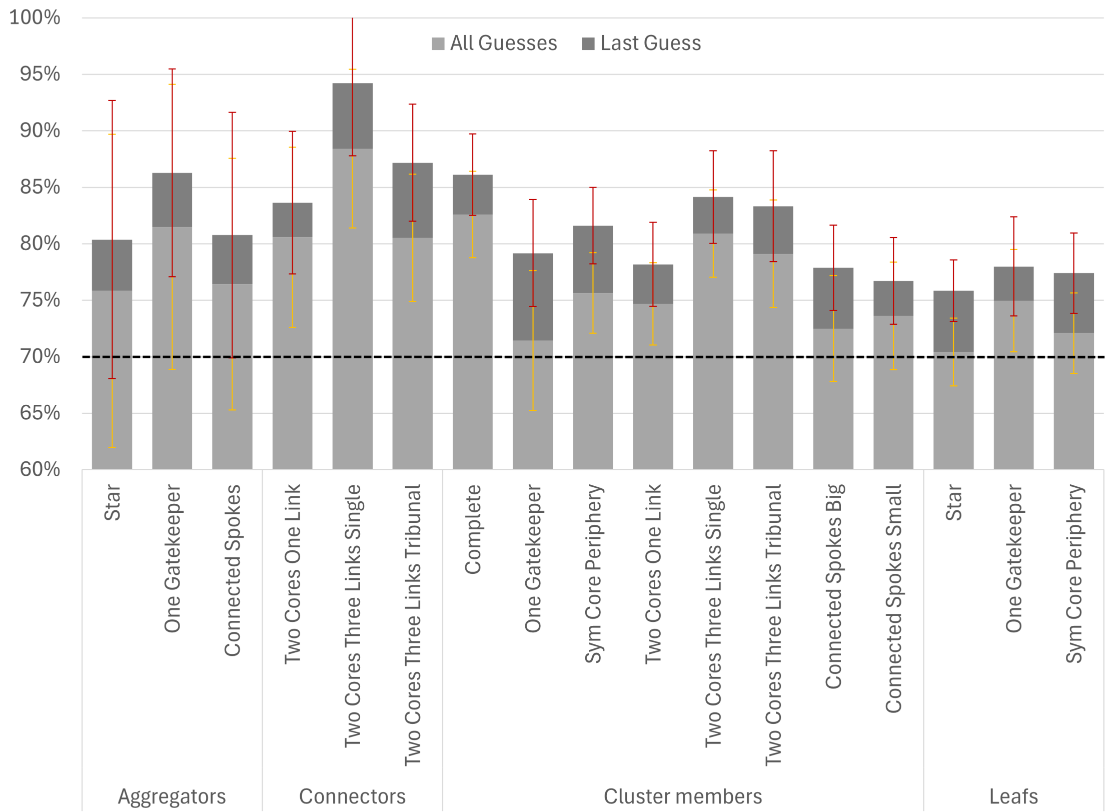


Figure 7: Overall and Long-term Performance, by Network Position

Notes: The frequency of correct guesses for all rounds in a game (gray bars) and for the final guess only (black bars). Whiskers denote 95% confidence intervals, with standard errors clustered at the participant level. The horizontal line at 70% indicates the probability of correct guess if one follows own signal and ignores everything else.

across positions and indicate which positions are most advantageous within and across networks.[58] Although large standard errors—driven by substantial variation across participants and signal distributions—limit precision, three key patterns emerge.

First, consider nodes connected to all others—namely, all agents in the Complete network and aggregators in *Single Aggregator* networks. By the end of the first round, these positions have access to nearly all private signals and should, in principle, achieve very high accuracy both overall and in the final round. In practice, however, *Under-Reaction to New Information* limits performance to below 87% correct guesses in both measures, which amounts to less than 60% of the expected improvement upon the benchmark of 70%.

Second, consider nodes that, by Proposition 1, should imitate an influencer—namely, non-aggregators in *Single Aggregator* networks, leafs in the Symmetric Core–Periphery network, and non-connectors in the Two Cores networks. Proposition 1 implies that these nodes should match their influencers' final-guess accuracy and approximate it on average. In reality, however, accuracy

[58]A companion paper Agranov et al. (2025) examines whether subjects' subjective perceptions align with these differences; see also Footnote 12.

rates for potential imitators are lower across all networks, positions, and measures, reflecting the *Under-Imitation* behavioral friction.

Third, degree alone does not have a consistent effect on performance. We observe a positive effect within networks when under-imitation leads higher-degree influencers to outperform potential imitators, and across networks when leafs generally perform worse than most other positions. Conversely, we find a negative effect within the One Gatekeeper network, where leafs outperform cluster members on average, and across networks, where connectors in the Two Cores networks achieve higher accuracy than aggregators in the Star and Connected Spokes networks, despite having lower degree. Taken together, these patterns show that degree alone cannot account for information aggregation performance; the broader network structure must be considered.

8 Discussion

8.1 The Experimental Literature on Information Aggregation on Networks[59]

Experimental research on small-group dynamics began in the 1950s with the "MIT Experiments" or "Bavelas Group Experiments" (see Shaw (1964) for a survey and follow-up work). In these studies, groups of 3–5 participants were placed in interconnected cubicles and communicated through written messages (or verbally) via wall slots (or intercom devices). Information about a puzzle was distributed among participants, who used the available communication network to collaboratively solve it.[60] These experiments demonstrated that the structure of the communication network significantly influenced problem-solving efficiency: centralized networks (e.g., star) outperformed others when information needed to be collected in a single location, while decentralized networks (e.g., cliques) were more effective when further processing was required. Central positions, however, often suffered performance declines under heavy cognitive load—a phenomenon described as "saturation," "vulnerability," or "over-information." For a modern counterpart, see Bernstein et al. (2023).

While in these puzzle-solving experiments, as in persuasion bias experiments,[61] each participant receives a unique piece of information essential to solving the task, information aggregation experiments differ in that each participant receives a noisy signal about the true state of the world. In this context, the network governs what information is available when forming beliefs about the external state.

Building on the experimental literature on social learning (e.g., Anderson and Holt (1997)),[62] Choi et al. (2005, 2012) were the first to study information aggregation over networks. Using directed networks of size 3 and signal accuracy $\frac{2}{3}$, they implemented a 3×3 design: three networks

[59] For an early survey see Section 2.5 in Choi et al. (2016).

[60] In the initial design, each participant received a card containing several symbols, with only one symbol common to all cards. The task was to identify this shared symbol (Bavelas (1950); Leavitt (1951)).

[61] In persuasion bias experiments, subjects receive noisy numerical signals and are incentivized to estimate the group average, relying on network-mediated information processing (Corazzini et al. (2012); Brandts et al. (2015); Battiston and Stanca (2015)).

[62] One-shot sequential learning designs on directed networks have been implemented in laboratory settings by Brown (2020), who use networks of size 5 with signal accuracy 0.7, and by Dasaratha and He (2021), who use random networks of size 40 with signals drawn from a Gaussian distribution.

(Complete, Star, Circle) crossed with three signal distribution conditions.[63] They found that a myopic Bayesian model fits the data well when augmented with exogenous logistic shocks to the preferences and allows subjects to respond to these trembles (Quantal Response Equilibrium model). Choi (2012) generalized this framework using a Cognitive Hierarchy QRE model and concluded that the dominant cognitive type is closely related to Bayes-rational behavior.

Subsequent work extended the analysis to slightly larger networks, where full Bayesian inference becomes cognitively demanding. Much of the recent literature therefore evaluates how well simpler alternatives—especially the naïve heuristic—describe observed behavior. Grimm and Mengel (2020) study undirected networks of size 7 and signal accuracy $\frac{4}{7}$, implementing a $3 \times 2 \times 3$ design: three network topologies (Star, Circle, Kite); two signal distributions (each with exactly four correct signals); and three information conditions.[64] They find that behavior varies across information treatments and interpret this as evidence against the naïve model. However, under full information, the naïve model performs comparably to Bayesian predictions at the aggregate level and outperforms it at the individual level. Moreover, they propose an adjusted naïve heuristic in which the weight on one's private signal increases with the clustering coefficient, while weights on neighbors' signals decrease and remain equal.[65] This adjusted rule outperforms both standard Bayesian and naïve models in two additional networks tested in the laboratory. Chandrasekhar et al. (2020) conduct two experiments—one with Indian villagers and another with Mexican university students—on undirected networks of size 7, using signal accuracy of $\frac{5}{7}$. They use predictions of the Bayesian and naïve models to find that the naïve model fits the behavior of Indian villagers significantly better than that of Mexican students. They then use structural estimation to fit a mixture of Bayesian and naïve agents in the experimental sample. Estimated shares of Bayesian agents are around 10% in the Indian sample and 50% in the Mexican sample.

A study of particular relevance to our setting is Choi et al. (2023), who investigate three directed networks of size 40 with signal accuracy 0.7. Their baseline Erdős–Rényi network is compared with a Stochastic Block model to study cohesiveness, and a pre-selected Royal Family network to explore hierarchy. Using primarily aggregate-level analysis, they conclude that the naïve model provides a much better fit than the myopic Bayesian benchmark.

[63] Full information: every player receives a signal; High information: every player receives a signal in probability $\frac{2}{3}$; Low information: every player receives a signal in probability $\frac{1}{3}$.

[64] No Information (only direct neighbors known), Incomplete Information (degree distribution known), and Complete Information (entire network revealed). In their 2-urn treatment, Mueller-Frank and Neri (2015) implemented a No Information condition and a similar design using networks of size 5 or 7 and signal accuracy $\frac{2}{3}$. They focus on testing behavioral axioms, which they later use to introduce a Quasi-Bayesian updating model.

[65] A different adjustment to the naïve heuristic is suggested by Jiang et al. (2023). They conduct a neuro-imaging study on undirected networks of size 7 with signal accuracy $\frac{5}{7}$, where one subject is scanned with fMRI while others participate from standard lab settings. Subjects observe neighbors' guesses sequentially (rather than simultaneously). They find that from the third round onward, brain activity reveals greater weight placed on guesses from well-connected neighbors.

8.2 Other Models of Information Aggregation

The central claim of this study is that the myopic Bayesian and naïve models cannot account for the behavior observed in our laboratory experiments. Our approach is to introduce, identify, and evaluate structural and behavioral frictions that align with patterns in the data. An alternative approach would be to adopt more sophisticated variants of the myopic Bayesian model or the naïve heuristic in an attempt to reconcile them with the experimental evidence while remaining within the same theoretical framework. In this subsection, we examine several such models and present evidence that they, too, are inconsistent with our findings.

We begin with myopic Bayesian models that introduce noise. A prominent example is the Quantal Response Equilibrium model of Choi et al. (2012), which assumes that agents follow a logit model of discrete choice.[66] This model further assumes that agents hold rational expectations regarding their neighbors' true error rates and use estimated error rates from the previous decisions to update their posterior beliefs. One indication of the inconsistency between noisy myopic Bayesian models and our data comes from second-round behavior: Table 5 shows that the rate of incorrect guesses in the second round among first-round minority members remains substantial, even when that minority is very small. Such behavior cannot be a rational response to first-round mistakes: only 7.8% of first-round guesses are incorrect, so the probability that small minorities in round 1 are correct is negligible. In addition, recall that in the spirit of Quantal Response Equilibrium model, we introduced a reduced-form model in Section 5.3 to assess whether *Under-Imitation* could be a rational response to *Under-Reaction to New Information*. When calibrated to the "true error rates" in our data, the model predicts imitation rates far higher than those actually observed.

We now turn to adjustments of the standard naïve model. As previously discussed, the standard naïve model fails to account for key patterns in our data: subjects often do not aggregate correctly in round 2 (Finding 7) and deviate from the model's predictions in round 3 (see Footnote 48). A common adjustment of the naïve model introduces unequal weighting, typically to reflect over-weighting of one's own signal. For instance, Grimm and Mengel (2020) observe that "relative to the naïve model, participants on average place too much weight on their own information." They propose a model in which the weight on a subject's own previous guess increases with its clustering coefficient—perhaps to account for correlation in the information received from neighbors—while the weights on neighbors' previous guesses remain equal. We can test two predictions of the rule they suggest. First, when the self-clustering coefficient is zero, this rule collapses to the myopic Bayesian model. Therefore, it predicts that the aggregator in the *Star* network—where the self-clustering coefficient is zero—should follow the myopic Bayesian prediction and aggregate correctly in round 2. Yet, our data show that this occurs in only 42% of cases where the aggregator is in the minority at the end of round 1. Second, a subject in the Complete network should switch less often than an aggregator in a *Single Aggregator* network of the same size, due to the higher weight assigned to their own guess. This prediction is contradicted by the evidence in Table 4. More generally, an

[66] An agent's random utility over alternatives depends on expected payoff and a private, standard Gumbel, idiosyncratic shock, i.i.d. across periods, agents, and actions.

interpretation of an unequal weighting rule—where the weights on neighbors' previous guesses are equal—is a fixed-threshold heuristic: a subject switches only if the number of neighbors with an opposing guess exceeds some fixed threshold. To test this, we elicited two subject-level measures: the maximum opposing majority size (MAX) for which they did not switch, and the minimum majority size (MIN) for which they did switch. We were able to get both numbers for 398 subjects. Only 65 of them (16.3%) satisfied the condition $MIN > MAX$, which is required for a fixed-threshold rule or a naïve model in which the subject assigns equal weights to all neighbors.

Next, consider a broader class of naïve models, which we refer to as the $\mathbf{w_i}(t)$ *heuristic*. In this framework, agents assign time-dependent weights to their own guess and to each of their neighbors' guesses in the previous period. These weights are fixed before the game begins and may vary across agents, neighbors, and rounds.[67] This general family encompasses the weights proposed by Grimm and Mengel (2020), since clustering coefficients can be computed ex ante. Crucially, however, all these heuristics assume static perceptions of neighbors: weights are fixed and unaffected by neighbors' observed behavior during the information aggregation process. This assumption contradicts our empirical findings on trust, discussed in Section 6.2. There, we show that imitation behavior depends on cues that increase confidence in the aggregator's guess—such as alignment with the local majority or evidence of responsiveness (e.g., the aggregator switching between rounds 1 and 2). These cues dynamically shape subjects' beliefs about the aggregator's credibility. Any $\mathbf{w_i}(t)$ heuristic is thus inconsistent with the observed trust-based imitation, as trust implies that different histories of play may lead to different weights—a feature these models explicitly rule out. Usually, the weakness of naïve updating is said to be its neglect of network structure, which Grimm and Mengel (2020) found to be inconsistent with their data. Our findings, however, highlight a distinct shortcoming: naïve models typically assume that agents' perceptions of their neighbors are fixed, unaffected by the evolving history of the game. Our evidence suggests that subjects do update their perceptions over time. In this sense, the naïve model simplifies complexity—by ignoring network structure—while also inadvertently eliminating a cognitively natural mechanism: adjusting perceptions of others based on experience over time. While one could imagine relaxing the model to allow weights to vary based on observed history, this might introduce too many degrees of freedom, undermining the model's explanatory power.[68]

Finally, Mueller-Frank (2014) and Chandrasekhar et al. (2020) propose models in which agents are either myopic Bayesians or naïve. In both models, a subject whose first-round guess differs from the local first-round majority is expected to switch in round 2. However, Table 4 shows that in

[67] Formally, let $B(i) = \{j_1, \ldots, j_b\}$ denote the set of direct neighbors of agent i. Her weight vector at time t is $\mathbf{w_i}(t) = (w_i^0(t), w_i^1(t), \ldots, w_i^b(t))$ where $\sum_{k=0}^{b} w_i^k(t) = 1$ and $\forall k \in \{0, \ldots, b\} : w_i^k(t) \geq 0$. $w_i^0(t)$ is the weight assigned to her own previous guess a_i^{t-1}, and $w_i^k(t)$ $(k \in \{1, \ldots, b\})$ is the weight on neighbor j_k's guess $a_{j_k}^{t-1}$. Under this heuristic, agent i calculates $\mathbb{1}_{a_i^{t-1}=W} w_i^0(t) + \sum_{k=1}^{b} \mathbb{1}_{a_{j_k}^{t-1}=W} w_i^k(t)$. She then sets $a_i^t = W$ if the calculation is greater than 0.5, or $a_i^t = B$ if it is less than 0.5; otherwise, she is indifferent. Choi et al. (2023) introduce a standard naïve model with trembles. Our deterministic version can easily be extended to include stochastic perturbations, but this extension does not add to the present discussion.

[68] Jansen (2024) proposes dynamic weighting between one's own guess and a weighted average of neighbors' guesses, in a different setting of information aggregation over networks.

at least 40% of such cases, subjects do not switch. Moreover, when both models predict that a potential imitator should switch in round 3, only 41.8% of subjects actually do so. These deviations indicate that a substantial share of the subject pool cannot be accurately classified as either myopic Bayesian or naïve.

8.3 Individual Decision Making over Networks

A natural next step is to understand the individual-level decision processes that produce *Under-Reaction to New Information* and *Under-Imitation*. This framework must also account for the behaviors observed when structural frictions arise, the evidence that imitation is trust-dependent, the contrast between behavior in the complete network and that of the aggregator in *Single Aggregator* networks, and other phenomena reported here.

One approach is to remain within the myopic Bayesian framework, augmented with components that explain our positional- and individual-level results.[69] Another approach is to remain within a heuristic framework but account more closely for the empirical findings: individuals exhibit switching aversion, they recognize positional differences, they adapt their perceptions dynamically, and they over-react to others' mistakes.[70] The simplicity of the chosen heuristic may depend on the network structure and complexity.[71] As in other areas of decision making, these directions are not mutually exclusive: better optimization models can inform our understanding of procedural thinking, and insights into procedural thinking can guide the development of better optimization models.

References

Acemoglu, D., Dahleh, M. A., Lobel, I., and Ozdaglar, A. (2011). Bayesian learning in social networks. *The Review of Economic Studies*, 78(4):1201–1236.

Acemoglu, D. and Ozdaglar, A. (2011). Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, 1(1):3–49.

Agranov, M., Gillen, B., and Persitz, D. (2024). A comment on "testing models of social learning on networks: Evidence from two experiments". *Econometrica*, 92(5):1–6.

Agranov, M., Gillen, B., and Persitz, D. (2025). When do we recognize the best information aggregator? Technical report, Working paper.

[69] Choi (2012) propose a cognitive hierarchy model in which subjects choose ex-ante the number of periods (τ) they will accumulate information from their neighbors. An agent who selects τ acts as a myopic Bayesian for the first τ periods, then persists with the optimal decision based on the information collected in the first $\tau - 1$ periods. Notably, Result 3 in Choi (2012) shows that incorporating cognitive hierarchy into the Quantal Response Equilibrium model requires a much smaller error rate to fit the experimental data in Choi et al. (2005, 2012).

[70] Grimm and Mengel (2020): "participants might be using rules of thumb that, although not Bayesian, are less naïve than the naïve model would suggest."

[71] See relevant discussions in Choi et al. (2005), Mobius et al. (2015), Kovářík et al. (2018), and Dasaratha and He (2021).

Agranov, M., Healy, P. J., and Nielsen, K. (2023). Stable randomisation. *The Economic Journal*, 133(655):2553–2579.

Aiello, J. R. and Douthitt, E. A. (2001). Social facilitation from triplett to electronic performance monitoring. *Group Dynamics: Theory, Research, and Practice*, 5(3):163.

Amelio, A. (2024). Social learning, behavioral biases and group outcomes. Working Paper.

American College of Surgeons (2024). Statement on bicycle safety and the promotion of bicycle helmet use. Accessed: 2025-02-28.

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

Augenblick, N., Lazarus, E., and Thaler, M. (2025). Overinference from weak signals and underinference from strong signals. *The Quarterly Journal of Economics*, 140(1):335–401.

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

Ba, C., Bohren, A., and Imas, A. (2024). Over- and underreaction to information. Working Paper.

Ba, C., Bohren, J. A., and Imas, A. (2025). The role of representational and computational complexity in belief formation. In *AEA Papers and Proceedings*, volume 115, pages 631–636. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Bala, V. and Goyal, S. (1998). Learning from neighbours. *The review of economic studies*, 65(3):595–621.

Battiston, P. and Stanca, L. (2015). Boundedly rational opinion dynamics in social networks: Does indegree matter? *Journal of Economic Behavior & Organization*, 119:400–421.

Bavelas, A. (1950). Communication patterns in task-oriented groups. *Journal of the acoustical society of America*.

Becker, J., Brackbill, D., and Centola, D. (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the national academy of sciences*, 114(26):E5070–E5076.

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

Berninghaus, S. K., Ehrhart, K.-M., and Keser, C. (2002). Conventions and local interaction structures: experimental evidence. *Games and Economic Behavior*, 39(2):177–205.

Bernstein, E. S., Shore, J. C., and Jang, A. J. (2023). Network centralization and collective adaptability to a shifting environment. *Organization Science*, 34(6):2064–2096.

Bigoni, M., Camera, G., and Gallo, E. (2025). Money and social exclusion in networks. Unpublished.

Bikhchandani, S., Hirshleifer, D., Tamuz, O., and Welch, I. (2024). Information cascades and social learning. *Journal of Economic Literature*, 62(3):1040–1093.

Blackwell, D. (1953). Equivalent comparisons of experiments. *The annals of mathematical statistics*, pages 265–272.

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.

Brown, C. (2020). Networks and economic behavior: Theory and experiments. Dissertation.

Cardoso, F. M., Gracia-Lázaro, C., Moisan, F., Goyal, S., Sánchez, Á., and Moreno, Y. (2020). Effect of network topology and node centrality on trading. *Scientific Reports*, 10(1):11113.

Cassar, A. (2007). Coordination and cooperation in local, random and small world networks: Experimental evidence. *Games and Economic Behavior*, 58(2):209–230.

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.

Charness, G., Feri, F., Meléndez-Jiménez, M. A., and Sutter, M. (2014). Experimental games on networks: Underpinnings of behavior and equilibrium selection. *Econometrica*, 82(5):1615–1670.

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*, 9:288–97.

Choi, S. (2012). A cognitive hierarchy model of learning in networks. *Review of Economic Design*, 16:215–250.

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

Choi, S., Gale, D., and Kariv, S. (2012). Social learning in networks: a quantal response equilibrium analysis of experimental data. *Review of Economic Design*, 16(2):135–157.

Choi, S., Galeotti, A., and Goyal, S. (2017). Trading in networks: theory and experiments. *Journal of the European Economic Association*, 15(4):784–817.

Choi, S., Gallo, E., and Kariv, S. (2016). Networks in the laboratory. In *The Oxford Handbook of the Economics of Networks*. Oxford University Press.

Choi, S., Goyal, S., Guo, F., and Moisan, F. (2024a). Experimental evidence on group size effects in network formation games. Working Paper.

Choi, S., Goyal, S., Guo, F., and Moisan, F. (2024b). Experimental evidence on the relation between network centrality and individual choice. Working Paper.

Choi, S., Goyal, S., Moisan, F., and To, Y. Y. T. (2023). Learning in networks: An experiment on large networks with real-world features. *Management Science*, 69(5):2778–2787.

Conlon, J., Mani, M., Rao, G., Ridley, M., and Schilbach, F. (2022). Not learning from others. Working Paper.

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.

Cottrell, N. B. (1972). Social facilitation. In McClintock, C. G., editor, *Experimental Social Psychology*, pages 185–236. Holt, Rinehart and Winston, New York.

Dasaratha, K. and He, K. (2021). An experiment on network density and sequential learning. *Games and Economic Behavior*, 128:182–192.

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

DeMarzo, P. M., Vayanos, D., and Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly journal of economics*, 118(3):909–968.

Enke, B. (2024). The cognitive turn in behavioral economics. Working Paper.

Esponda, I., Oprea, R., and Yuksel, S. (2023). Seeing what is representative. *The Quarterly Journal of Economics*, 138(4):2607–2657.

Gale, D. and Kariv, S. (2003). Bayesian learning in social networks. *Games and economic behavior*, 45(2):329–346.

Geanakoplos, J. D. and Polemarchakis, H. M. (1982). We can't disagree forever. *Journal of Economic theory*, 28(1):192–200.

Gneezy, U. and Potters, J. (1997). An experiment on risk taking and evaluation periods. *The quarterly journal of economics*, 112(2):631–645.

Golub, B. and Jackson, M. O. (2010). Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–149.

Golub, B. and Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, 127(3):1287–1338.

Golub, B. and Sadler, E. (2016). Learning in social networks. In Bramoullé, Y., Galeotti, A., and Rogers, B., editors, *The Oxford Handbook of the Economics of Networks*. Oxford University Press.

Goyal, S. and Vega-Redondo, F. (2005). Network formation and social coordination. *Games and Economic Behavior*, 50(2):178–207.

Grant, D. A., Hake, H. W., and Hornseth, J. P. (1951). Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement. *Journal of experimental psychology*, 42(1):1–5.

Green, J. and Stokey, N. (1978). Two representations of information structures and their comparisons. Working Paper.

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

Guerin, B. (2010). Social facilitation. In *The Corsini encyclopedia of psychology*. Wiley Online Library.

Humphreys, L. G. (1939). Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *Journal of Experimental Psychology*, 25(3):294–301.

Jackson, M. O. (2008). *Social and Economic Networks*. Princeton University Press.

Jackson, M. O. (2019). The friendship paradox and systematic biases in perceptions and social norms. *Journal of political economy*, 127(2):777–818.

Jackson, M. O. and Rogers, B. W. (2005). The economics of small worlds. *Journal of the European Economic Association*, 3(2-3):617–627.

Jackson, M. O. and Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, 71:44–74.

Jansen, Y. (2024). Social learning and stubbornness. Working Paper.

Jiang, Y., Mi, Q., and Zhu, L. (2023). Neurocomputational mechanism of real-time distributed learning on social networks. *Nature neuroscience*, 26(3):506–516.

Kearns, M. (2012). Experiments in social computation. *Communications of the ACM*, 55(10):56–67.

Kirchkamp, O. and Nagel, R. (2007). Naive learning and cooperation in network experiments. *Games and Economic Behavior*, 58(2):269–292.

Kovářík, J., Mengel, F., and Romero, J. G. (2018). Learning in network games. *Quantitative Economics*, 9(1):85–139.

Lajunen, T. and Räsänen, M. (2004). Can social psychological models be used to promote bicycle helmet use among teenagers? a comparison of the health belief model, theory of planned behavior and the locus of control. *Journal of safety research*, 35(1):115–123.

Leavitt, H. J. (1951). Some effects of certain communication patterns on group performance. *The journal of abnormal and social psychology*, 46(1):38.

Lee, L. K., Flaherty, M. R., Blanchard, A. M., and Agarwal, M. (2022). Helmet use in preventing head injuries in bicycling, snow sports, and other recreational activities and sports. *Pediatrics*, 150(3):e2022058878.

Lerman, K., Yan, X., and Wu, X.-Z. (2016). The" majority illusion" in social networks. *PloS one*, 11(2):e0147617.

Loomes, G. (1998). Probabilities vs Money: A Test of some Fundamental Assumptions About Rational Decision Making. *The Economic Journal*, 108(447):477–489.

McFarland, D. A., Moody, J., Diehl, D., Smith, J. A., and Thomas, R. J. (2014). Network ecology and adolescent social structure. *American sociological review*, 79(6):1088–1121.

Mobius, M., Phan, T., and Szeidl, A. (2015). Treasure hunt: Social learning in the field. Working Paper.

Mobius, M. and Rosenblat, T. (2014). Social learning in economics. *Annu. Rev. Econ.*, 6(1):827–847.

Molina-García, J., Queralt, A., Bengoechea, E. G., Moore, A., and Mandic, S. (2018). Would new zealand adolescents cycle to school more if allowed to cycle without a helmet? *Journal of Transport & Health*, 11:64–72.

Morris, S. (2000). Contagion. *The Review of Economic Studies*, 67(1):57–78.

Mossel, E., Sly, A., and Tamuz, O. (2015). Strategic learning and the topology of social networks. *Econometrica*, 83(5):1755–1794.

Mueller-Frank, M. (2013). A general framework for rational learning in social networks. *Theoretical Economics*, 8(1):1–40.

Mueller-Frank, M. (2014). Does one bayesian make a difference? *Journal of Economic Theory*, 154:423–452.

Mueller-Frank, M. and Neri, C. (2015). A general model of boundedly rational observational learning: Theory and evidence. Unpublished.

Mueller-Frank, M. and Neri, C. (2021). A general analysis of boundedly rational learning in social networks. *Theoretical Economics*, 16(1):317–357.

Persitz, D. (2010). Core-periphery r&d collaboration networks. Technical report, Working paper.

Rigotti, L., Wilson, A., and Gupta, N. (2023). The experimenters' dilemma: Inferential preferences over populations. Unpublished.

Rivas, J. (2013). Probability matching and reinforcement learning. *Journal of Mathematical Economics*, 49(1):17–21.

Rubinstein, A. (2002). Irrational diversification in multiple decision problems. *European Economic Review*, 46(8):1369–1378.

Schelling, T. C. (1973). Hockey helmets, concealed weapons, and daylight saving: A study of binary choices with externalities. *Journal of Conflict resolution*, 17(3):381–428.

Shaw, M. E. (1964). Communication networks. In *Advances in experimental social psychology*, volume 1, pages 111–147. Elsevier.

Siegel, S. and Goldstein, D. A. (1959). Decision-making behavior in a two-choice uncertain outcome situation. *Journal of Experimental Psychology*, 57(1):37–42.

Verlinde, L., Verlinde, F., Van Doren, S., De Coninck, D., and Toelen, J. (2024). Cycle safe or cycle cool? adolescents' views on bicycle helmet use and injury prevention campaigns in belgium. *Injury prevention*.

Weizsäcker, G. (2010). Do we follow others when we should? a simple test of rational expectations. *American Economic Review*, 100(5):2340–2360.

Woodford, M. (2020). Modeling imprecision in perception, valuation, and choice. *Annual Review of Economics*.

Ziegelmeyer, A., March, C., and Krügel, S. (2013). Do we follow others when we should? a simple test of rational expectations: comment. *American Economic Review*, 103(6):2633–2642.