Competing for Influence in Networks Through Strategic Targeting

Margherita Comola* Agnieszka Rusinowska§ Marie Claire Villeval¶

April 8, 2026

Abstract

We experimentally investigate targeting decisions in a setting where a human player competes for influence in a network against a computerized opponent with opposing views, whose targeting choice is revealed before the player acts. By varying network structure, opponent influence, and nodes opinion heterogeneity, we find that players typically adopt best-response strategies based on relative influence. However, they sometimes deviate – for example, by erroneously targeting central nodes or by avoiding the opponent's target. Targeting is also affected by affinity and opposition biases, the strength of which depends on the initial opinion distribution. Targeting the center, avoiding the competitor's target, or selecting nodes based on their initial opinions when these are not best responses generates significant efficiency losses.

Keywords: Network; Influence; Targeting; Competition; Laboratory Experiment

JEL codes: C91, D85, D91

Acknowledgements: We are grateful to F. Moisan and to participants at the SAET conference in Santiago, the SING conference in Messina, the CTN Workshop in Budapest, the Economic Science Association World meetings in Lyon, the ASFEE meeting in Grenoble, the European Conference on Networks at the University of Essex, and seminar presentations at Paris School of Economics and Hamburg University for useful comments. We are thankful to Q. Thevenet for programming the experiment. This research has benefited from the financial support of the French National Research Agency (ANR-21-CE26-0002-01 and ANR-17-EURE0001), IDEXLYON from Université

*Université Paris-Saclay and Paris School of Economics. Email: margherita.comola@psemail.eu

§CES, CNRS - University Paris 1 Panthéon-Sorbonne and Paris School of Economics. Email: Agnieszka.Rusinowska@univ-paris1.fr

¶CNRS, Université Lumiére Lyon 2, Université Jean-Monnet Saint-Etienne, emlyon business school, GATE, 69007, Lyon, France. Email: villeval@gate.cnrs.fr

de Lyon (INDEPTH-IDEX/SBP/2018/03) within the Programme Investissements d'Avenir (ANR-16-IDEX-0005) operated by the French National Research Agency, and the LABEX CORTEX (ANR-11-LABX-0042).

1 Introduction

Identifying the optimal target in a social network to spread information or secure a strategic advantage is a fundamental challenge in many contexts where actors compete for influence, whether brands seeking key opinion leaders to endorse their products, political campaigns engaging community leaders, companies forming strategic alliances with key partners, organizations attempting to shape information flows online, or cybersecurity services attacking and defending critical nodes in a network. Despite its empirical relevance and the extensive literature on identifying optimal targets and key players in social networks (for economic surveys, see, *e.g.*, Jackson, 2008; Bloch, 2016; Bramoullé et al., 2016; Zenou, 2016),[1,2] much less is known about how competition between influencers affects target selection. Existing empirical work overwhelmingly considers settings in which an actor has a monopoly on targeting, leaving open how individuals behave when a rival attempts to sway the same population.

Our study examines how individuals choose whom to target to maximize the spread of their opinions within a network when they face an opponent with opposing views, and it identifies the potential biases that arise in this process. In the context of information diffusion in networks, our study draws from the theoretical insights by Grabisch et al. (2018) in which two strategic agents with opposing fixed opinions engage in a constant-sum targeting game. Both agents choose a node to target simultaneously, and non-strategic nodes update their opinions according to the DeGroot process (DeGroot, 1974). Each non-strategic agent repeatedly revises their opinion by taking a weighted average of their neighbors' current opinions. This updating

[1]In economics, the theoretical literature on network targeting was pioneered by Ballester et al. (2006) who have inspired theoretical investigations (*e.g.*, König et al., 2014; de Marti and Zenou, 2015; Bloch and Shabayek, 2023), as well as empirical studies concerning a variety of issues, such as crime (Ballester et al., 2010; Liu et al., 2012), education (Calvó-Armengol et al., 2009; Hahn et al., 2015), R&D (König et al., 2014), technology adoption (Beaman et al., 2021), finance (Battiston et al., 2012; Cabrales et al., 2016; Demange, 2018) and micro-finance (Banerjee et al., 2013).

[2]Targeting in networks has also been widely studied outside economics. For example, computer sciences and complex systems literature study target selection for the optimal diffusion of new products (*e.g.*, Goldenberg et al., 2001), using an algorithmic perspective (*e.g.*, Domingos and Richardson, 2001; Richardson and Domingos, 2002; Kempe et al., 2003, 2005).

process occurs iteratively across the network until opinions converge to a stable equilibrium. This environment involves strategic uncertainty: because targeting occurs simultaneously, each strategic agent must form beliefs about the opponent's strategy, and equilibrium behavior depends on those beliefs. Three fundamental elements jointly determine optimal targeting: i) the opponent's chosen target (*i.e.*, a central *vs.* a peripheral node), ii) the relative influence power of the opponent (*i.e.*, how the weight placed on the opponent's opinion compares with the influence exerted by other sources on the connected nodes), and iii) the network's topology. For example, copying an opponent targeting a central node may be optimal when both players have similar influence; but when the opponent is substantially more influential, steering the network may instead require targeting a peripheral –more influenceable– node. The point at which this strategic switch occurs can vary with the network's topology.

We take this model to the laboratory in a simplified environment designed to isolate best-response behavior. Rather than the simultaneous-move game in Grabisch et al. (2018), we implemented a sequential game in which the opponent's action and relative influence are known. We used small networks to limit cognitive load while still maintaining variation in the optimal targeting strategy. An automated opponent (with a fixed opinion) would first choose a node to target. The human participant, holding the opposite opinion, then observed the opponent's choice and influence level, and selected their own target in order to pull the network's average opinion toward their side. Because our primary interest lies in the targeting decisions rather than the opinion-updating process, the nodes consisted of automated players. Information spread throughout the network, and nodes' opinions were updated using the DeGroot process.[3] This setup created a constant-sum game.

This sequential and deterministic environment has two advantages. First, it removes ambiguity, risk, and belief formation about the opponent's behavior: the opponent's behavior is known rather than anticipated. Second, because the mapping from the pair of targets to final payoffs is deterministic and transparent, optimal

[3]Some experimental evidence shows that the DeGroot rule is not only simple but also often accounts for observed human behavior in network experiments (*e.g.*, Corazzini et al., 2012; Chandrasekhar et al., 2020).

play is objectively defined. Thus, any deviation from best-response choice can be attributed to decision errors and heuristics rather than strategic uncertainty. This approach allowed us to focus on a narrower, but analytically important, question: whether participants best respond when the strategic environment is fully transparent. We view this simplification as an imperfect but valuable step in understanding targeting competition. In a simultaneous environment, many forces can drive deviations from equilibrium predictions: ambiguity and risk aversion, biased beliefs about others' strategies, or mistakes in computing optimal responses. By eliminating uncertainty, our design isolates failures to best respond in a setting where optimal play is objectively defined and less complex. Estimating the frequency and determinants of failures to best respond in this simpler setting provides a benchmark for understanding behavior in targeting competition under strategic uncertainty.

To generate interesting variations in best responses, we manipulated the three key drivers highlighted by the theory. We varied the opponent's relative influence (the degree to which their opinion is listened to by others) while holding the participant's influence fixed, creating situations in which the best response shifts between targeting a central or a peripheral node. In the Increasing Influence treatment, the two targeting players started with the same degree of influence but the opponent's influence increased exogenously over the periods. We also varied the network structure by using four networks (a line, a kite, a star, a butterfly). Finally, we varied the opponent's target choice (center *vs.* periphery). In total, participants played eight blocks of five periods each: within a block, the network structure was kept constant, and the opponent's degree of influence increased across periods. For each network configuration, participants played one block where the opponent targeted the center and another block where the opponent targeted the periphery. Both types of opponent's targets were observed across all degrees of influence.

We also exploited a feature of the DeGroot framework: because the two targeting players are stubborn (*i.e.*, uninfluenced by others' opinions), the initial opinions of the nodes do not affect the distribution of final opinions or payoffs; they only determine how quickly the network reaches the steady state, not where it ends up.[4] This

[4]We acknowledge that it also has limitations. Since the final equilibrium does not depend on

makes initial opinions theoretically irrelevant for best responses, but we anticipated that in practice, players might still base their decisions on these initial opinions. We therefore manipulated the nodes' original opinions. In half of the blocks, all nodes started neutral (*i.e.*, they were indifferent between the two targeting players), while in the other half, they were randomly assigned heterogeneous initial opinions. Introducing a diversity of opinions allowed us to assess whether targeting players exhibit affinity or opposition biases, preferring to target nodes whose initial opinions are closer to their own, or avoiding nodes closer to the opponent, despite this being payoff-irrelevant.

For robustness, we further included two between-subjects control conditions, only for methodological reasons. In the Decreasing Influence treatment, we reversed the order in which influence changes across periods. In the Layout treatment, we altered the visual display of the networks on the screen to prevent the central node from appearing in the center of the screen. The findings of the main treatment hold in these controls.

Our design thus allows us to ask the following questions. To what extent do individuals adopt best-response strategies in a fully transparent and deterministic targeting game? When they deviate from optimal play, do they display systematic heuristics such as preferring central nodes or avoiding the opponent's target? Do they exhibit biases with respect to initial opinions that should be strategically irrelevant? Answering these questions allows us to explore how large the efficiency losses are when players choose suboptimal targets in the network.[5] In doing so, we complement recent work showing that heuristically guided targeting can perform nearly as well as optimal network-based strategies (*e.g.*, Akbarpour et al., 2025).

Our study yields three key findings. First, the best response was selected approximately 70% of the time (75% when the best response was targeting the center

the initial opinion of the non-strategic nodes, this model is not designed to capture the dynamics of opinion polarization fully.

[5]Note that, even in a simplified form, the multi-agent setting uniquely positions us to explore these factors, as the preference for targeting the center of a network, for example, cannot be tested in the analogous single-agent setting, as the game would always converge to the same steady state regardless of the targeting choice.

and 59% when it was targeting the periphery), with a sensitivity of choices to the complexity of decision-making, which depends on the combination of the network's structure, the opponent's level of influence, and the position in the network of the opponent's selected target. In particular, the best response rate was higher in star and butterfly networks, that is, in networks with a clear focal point or multiple best responses. Higher success was also observed among participants with stronger mathematical and strategic reasoning abilities and with experience over time.

Second, we uncovered behavioral regularities that depart from equilibrium strategies: conditional on best responses, participants exhibited a residual tendency to target nodes in the center and to avoid nodes targeted by their opponent even when this deviates from best-response strategies, consistent with a preference for visibility and differentiation.

Third, participants showed affinity and opposition biases: they were more likely to target nodes whose initial opinions are closer to their own, and displayed a (moderate) aversion to nodes whose initial opinions were closer to their opponent's, despite initial opinions having no impact on final payoffs. These biases were less pronounced in a balanced setting where different opinions are represented by multiple nodes, and they were highest when the opinion closer to the player's opinion is represented by one node only.

These findings highlight the complexity of targeting behavior in networks and reveal substantial monetary losses from failing to target the right individuals, even in a sequential environment without strategic uncertainty. Targeting a central node when it was not a best response led to an average 13.3% loss, while targeting a peripheral node not chosen by the opponent resulted in a 9.9% loss. In the Heterogeneous condition, average earnings were significantly lower when the best response was to target a node whose opinion was most distant from the player's. Such errors would be likely more frequent when information about the network structure is incomplete.

2 Related literature

Our research connects to three strands of network literature: studies on targeting, scompetition in targeting, and experiments on networks. It relates to the literature on diffusion (Jackson and Yariv, 2011; Lamberson, 2016), learning (Golub and Sadler, 2016), and contagion (Cabrales et al., 2016) in networks. A central question is how to identify optimal targets for information diffusion. Galeotti and Goyal (2009) show that optimal influence strategies depend on the mode of interaction, leading to targeting either low- or high-connectivity individuals. In Chatterjee and Dutta (2016), a firm strategically places an implant to propagate its product, choosing nodes that maximize decay centrality for high-quality products or nodes with high degree for low-quality ones. Focusing on optimal contracting, Belhaj and Deroïan (2019) show that a principal benefits from engaging only a subset of agents, avoiding central nodes when interaction intensity is high (see also Belhaj et al., 2023). Tsakas (2017) models a planner's optimal targeting to maximize behavioral diffusion in environments where agents imitate successful peers. Studying how the position of the first recipient shapes diffusion, Banerjee et al. (2013) model a word-of-mouth diffusion with an application to microfinance data. Diffusion centrality measures how extensively information spreads from a node, and understanding transmission is key for identifying optimal injection nodes.[6] In this context, Akbarpour et al. (2025) show that randomly selecting a small additional set of individuals can lead to more effective diffusion than targeting the optimal number of individuals over the network structure.

Our study distinguishes itself by focusing on target competition and adopting an empirical perspective, an area that few studies address.[7] Among the exceptions, Bimpikis et al. (2016) model firms' targeted advertising in a competitive setting,

[6]Other studies analyze targeting in politics, notably in the context of vote buying, revealing a preference for targeting reciprocal and well-connected citizens (*e.g.*, Finan and Schechter, 2012; Fafchamps and Labonne, 2020; Ravanilla et al., 2022), citizens with a weak ideological attachment (Dixit and Londregan, 1996), core supporters (Nichter, 2008), or depending on their political attitude and position in a network (Duarte et al., 2025).

[7]Studies on oligopolistic or perfect competition in networks usually focus on pricing strategies, rather than targeting influence (*e.g.*, Banerji and Dutta, 2009).

showing that marketing budgets are inefficiently high at equilibrium, with inefficiency increasing with the absorption centrality of targeted agents. Goyal et al. (2019) model contagion between firms using budgets to seed consumers in a network for product adoption. Our approach differs by modeling target competition in terms of opinion diffusion, focusing on how strategic agents take into account their competitor's relative influence, the network structure, and initial opinions. Targets are characterized by influenceability and intermediacy centrality (Grabisch et al., 2018), which differs from other centrality measures by accounting for opponents' targets.

Our study shares some features with the literature on Colonel Blotto games (*e.g.*, Roberson, 2006; Chowdhury et al., 2013; Kovenock and Roberson, 2021). In these zero-sum games, two competing players allocate limited resources across multiple battlefields. Players face a trade-off: concentrate resources on a few battlefields or spread them more widely. While we share with these games an interest in the strategic decision of where to exert influence under competitive pressure, our focus is instead on networks with interconnected nodes. Our contribution lies in studying whether players can identify a single entry point that maximizes the dissemination of their opinion within the network, relative to their opponent's.

Adopting an experimental approach allows us to identify the behavioral patterns that lead to deviations from best response strategies. Thus, we also contribute to the relatively scarce experimental literature on social networks (see surveys by Breza, 2016; Choi et al., 2016). Most experimental economics studies focus on whether individuals play the equilibrium in network games but do not consider targeting,[8] while experiments in psychology have shown that targeting is influenced by factors like personality, position in the network, or initial opinions.[9]

[8]Charness et al. (2014) explored the role of incomplete information in equilibrium selection in various networks, testing predictions from Galeotti et al. (2010) in games of strategic complements and substitutes. Rosenkranz and Weitzel (2012) used a public goods game to test the model by Bramoullé and Kranton (2007) with strategic substitutes. Gallo and Yan (2023) tested the game by Ballester et al. (2006) with strategic complements across a large strategy space to investigate the establishment of a social norm, showing the importance of the position in a network. Friedman et al. (2024) use coordination games to study social influence in local information networks, examining how an individual's position affects their ability to influence others' decisions directly or indirectly.

[9]Using fictitious networks, Smith and Carpenter (2018) found that subjects with strong intentions to buy antibiotic-free food tend to target nodes with high eigenvector and closeness centralities,

3 Experimental design and procedures

3.1 Design

The Targeting Game. As in the model of Grabisch et al. (2018) (summarized in Appendix A), two agents, called *Player* and *Opponent*, interact with n non-strategic *nodes* (n=5). These nodes are connected through an exogenous and undirected network. They have an initial opinion on a matter of interest in the interval $[0, 1]$.[10] They update their opinion by interacting repeatedly along network lines. This learning mechanism is mechanical as in the DeGroot model: the opinion of each node is determined by the average of the opinions of the nodes they are connected to.

Player and *Opponent* have fixed and opposite opinions on a matter of interest: Player has opinion 1, and Opponent has opinion 0. Their aim in this constant-sum game is to steer the average opinion of the network towards their own. They maximize their payoff by minimizing the distance between their own opinion and the average opinion of the non-strategic agents. To do so, each of them targets one node in the network to create a link with. Their strategy set is the set of all nodes.

The impact of the strategic agents on the nodes' opinions depends on an exogenous influence factor representing their *social* importance. The influence factor dictates how much weight their opinions carry when influencing a node. The Player's degree of influence is fixed and standardized to 1, that is the Player's opinion has the same weight as the opinion of any node connected to the target node. In contrast, the Opponent's degree of influence (denoted λ) varies. $\lambda = X$ means that the Opponent's opinion counts X times more for the targeted node than the opinion of each of the nodes the targeted node is linked to. For example, if $\lambda = 1$, the Opponent's opinion is equally as important as that of any other node or the Player; if $\lambda = 2$, the Opponent's opinion is twice as important as that of a standard node or the Player. In other words, λ scales the weight of the Opponent's influence. For a targeted node,

and those with higher centrality themselves target more central nodes. Bechler et al. (2020) found a higher inclination to persuade people with initially negative opinions to shift toward a positive one rather than attempting to make those with already positive opinions even more positive.

[10] In the experiment, we did not specify the nature of the topic. However, to make it less abstract, we indicated to the participants that they could think of a political, economic, or social topic.

the higher the value of λ, the more heavily the node is influenced by the Opponent's opinion compared to others in the network.[11]

Once the Player and the Opponent have chosen their targets, the updating mechanism of nodes is a weighted average of opinions and converges (*i.e.*, produces a stable opinion vector) after a number of iterations. Note that while the DeGroot updating mechanism is simultaneous, in our setting targeting choices are sequential (*i.e.*, the Player is informed of the Opponent's choice before picking his or her target). This sequential setup enables the study of players' best responses and reflects real-world scenarios, where targeting decisions often respond to others' prior moves. The updating process and the fact that the opinion revision process continues until all the opinions stabilize in the network (not necessarily on the same value) was explained in detail in the instructions. Moreover, in the practice and during the game, participants could visually observe the evolution of each node's opinions in real time on their screen and the final opinions in the network after the steady state was reached.

In the experiment, the five nodes and the Opponent were computerized agents, while the Player was a human subject.[12] This setup and the rule for opinion updating in the network were made common knowledge. Participants earned 600 ECU (Experimental Currency Units, with 110 ECU=€1) if they minimized the distance between the nodes' average opinion and their own opinion given the network and λ. Specifically, the actual reward was based on the share of the maximum attainable payoff and computed as (final average opinion / maximum average opinion attainable) x 600. Thus, a best-responding player always got a payoff of 600 ECU, which sets a natural metric for efficiency, as explained in Section 4.[13]

[11]In the instructions (see Appendix B.1), we presented λ as "the degree of influence", and we showed a simplified example with only three Bs and one A, and a figure. In this example, before A created a link, the weight of B2's opinion on B3 was 1. After A created a link with B3, if the degree of influence of A was 1, the weight of the opinion of B2 on B3 became 0.5 and that of the opinion of A on B3 was 0.5. If the degree of influence of A was 3, the weight of the opinion of B2 on B3 became 0.25 and that of the opinion of A on B3 was 0.75.

[12]Using computerized agents avoided prosocial considerations and strategic uncertainty that could affect Players' decisions. It facilitated identifying the players' learning of best response strategies.

[13]Note that when the opinion of the Opponent is preponderant, the player has limited scope for action. That is, as λ increases, the maximum average opinion of nodes attainable decreases, even if she best responds. If we were to remunerate subjects based only on the nodes' average final opinion,

Variations. In what follows, we describe our main treatment, named the Increasing Influence treatment. Figure 1 represents the within-subject variations in panel A, and within- and between-subject variations in panel B. Each session consisted of eight blocks of five periods each, where participants made a targeting choice, providing 40 observations per participant. Within each block, we varied the Opponent's degree of influence λ from 1 to 5 in a fixed ascending order over the five periods.[14]

Between blocks, we varied within subjects the network configuration and the Opponent's targeting choice between center and periphery (Figure 1, panel A), while keeping these dimensions constant within each block. We selected four network configurations with well-behaved properties (*i.e.*, strongly connected) and meaningful visual structures: a line, a kite, a star, and a butterfly. Each configuration appeared in two blocks, one in which the Opponent targeted the center and another one in which it targeted the periphery. The order of blocks was random at the subject level.

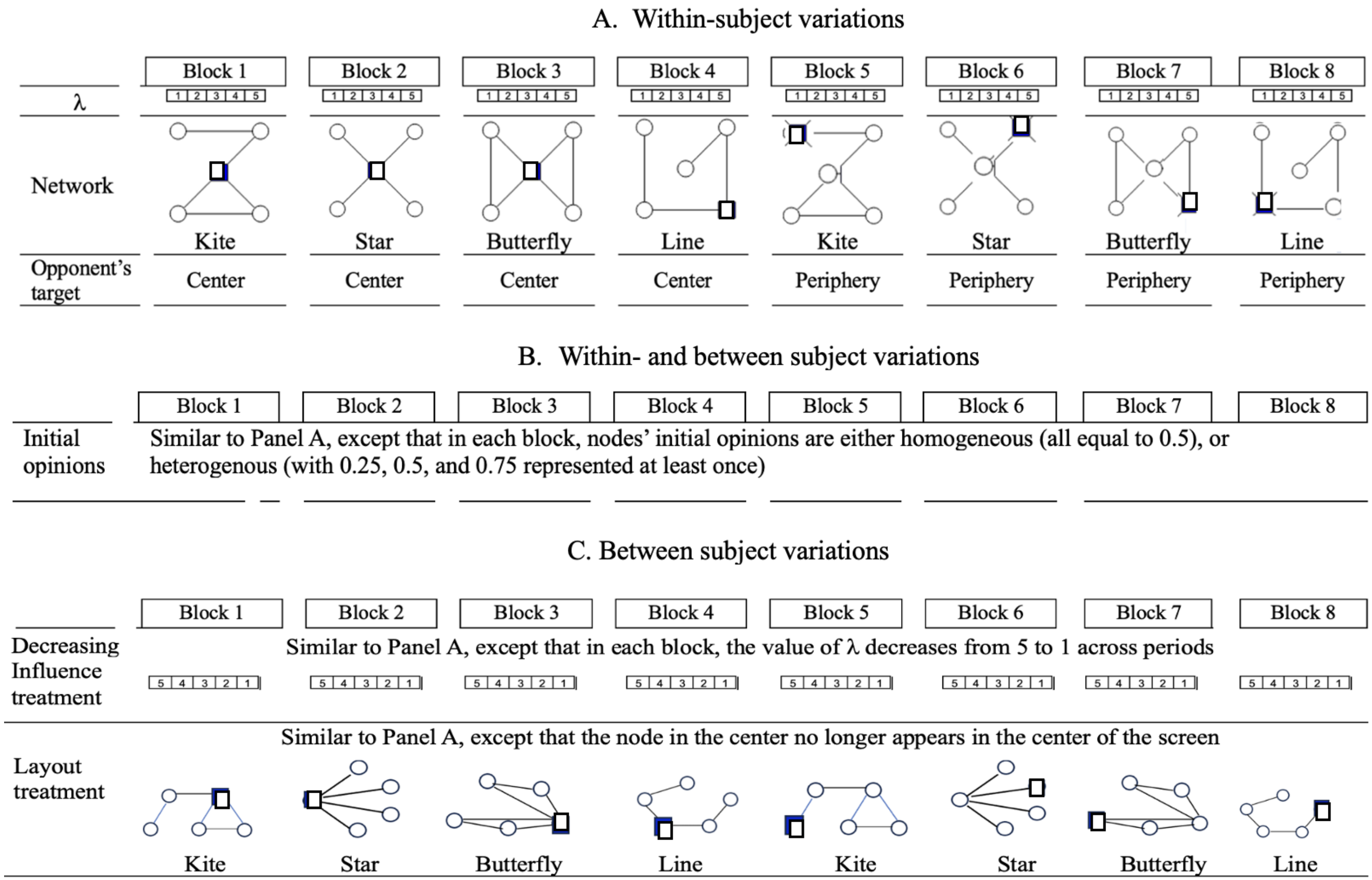
We also manipulated, within and between subjects, the nodes' initial opinions (Figure 1, panel B). Indeed, we anticipated that in an influence competition, players could be sensitive to opinion proximity with a node, although theoretically irrelevant. Subjects faced in random order four blocks where all nodes' initial opinion was equal to 0.5 (*"homogeneous blocks"*) and four other blocks where the nodes were assigned different initial opinions (0.25, 0.5, and 0.75) (*"heterogeneous blocks"*). In the latter, each node was assigned one of the three possible opinions through randomization, under the constraint that all three values had to be represented at least once across the five nodes. Because of this constraint, the number of nodes assigned to any given opinion ranged from one to three in each network. The distribution of initial opinions was node- and block-specific, and initial opinions were fixed within a block. In both homogeneous and heterogeneous blocks, the Opponent targeted the center in two of them and it targeted the periphery in the two others. Each network structure appeared in one homogeneous block and one heterogeneous block.

---

we would not take this into account. By using a payment rule based on the maximum attainable payoff, we created a common incentive that does not depend on the variations in the protocol.

[14]We chose this ascending order, rather than a randomized order across periods, to facilitate participants' learning.

Figure 1: Outline of the experimental design (part 2)



Notes: Panel A illustrates the within-subject variations in the Increasing Influence treatment. Players played eight blocks in random order. Within each block, they played 5 periods with an increasing degree of influence of the Opponent over periods (λ). Each network structure appeared twice (random order), once with the Opponent (depicted by a square) targeting the center, and once with the Opponent targeting a peripheral node. Panel B shows within-and between-subject variations. In the four homogeneous blocks, all nodes had the same initial opinion. In the four heterogeneous blocks, each of the three possible initial opinions was represented at least once.

To ensure participants understood that their payoff was based on the final opinions in the steady state, rather than on the initial opinions, they practiced the task for 3 minutes in a homogeneous setting and 3 minutes in a heterogeneous setting, and they could visually follow the color-coded process of convergence of opinions.

All the rules were common information in the instructions. At the beginning of a block, participants were informed about the network structure and the Opponent's targeting choice in the block. At the beginning of each period, they were reminded of the Opponent's degree of influence λ and each node's initial opinion.

The participants' screen displayed a visual representation of the network, showing the position and the initial opinion of each node next to it, and a white avatar representing the Opponent next to the node it targeted (see Figure C1 in online Appendix C). Once the participant made a decision, a black avatar representing this participant appeared next to the targeted node. At the beginning of each period, nodes were color-coded on a scale from white to black based on their initial opinion (*e.g.*, gray representing the opinion 0.5). As the opinion spread through the network, the nodes turned darker or whiter depending on the two targeting decisions. Each node's opinion and the color of the nodes changed continuously until convergence. At the end of each period, participants observed the final opinion of each node and received feedback on the average final opinion in the network and their payoff.

Control conditions. Two features of this design could influence behavior in the experiment, although irrelevant in the theoretical model. First, the Opponent's degree of influence always increased across periods. Second, the center of the network was always displayed in the center of the computer screens, which could induce a tendency to target the center. Therefore, we added *ex-post* two between-subject control conditions for purely methodological reasons (see panel C in Figure 1).

The only difference with the previously described *Increasing Influence treatment* is that in the *Decreasing Influence treatment*, the Opponent's degree of influence λ decreased from 5 to 1 in a fixed descending order over the five periods within each block. Comparing behavior between these two treatments helps determine whether individuals are sensitive to the direction in which their relative influence changes.

In the *Layout treatment*, the only difference with the *Increasing Influence treatment* is that we altered the graphical layout so that the central node was no longer depicted in the center of the graph (see panel C). Comparing behavior between these two treatments helps determine whether a bias towards centrality is induced by the experimental design itself or by the participants' inherent tendency.

Elicitation of strategic reasoning and cognitive abilities. Because these skills may influence participants' ability to best respond in the Targeting Game, we measured strategic reasoning and cognitive abilities. In part 1, participants played a Beauty Contest Game to assess their depth of strategic reasoning (see, *e.g.*, Nagel, 1995). They chose a number between 0 and 100, with a €10 prize for being closest to two-thirds of the session average. Lower numbers indicate deeper strategic reasoning.[15] In part 3, participants completed six Raven matrices in six minutes, earning ECU50 (€0.45) per correct answer, which provided a measure of cognitive abilities.

3.2 Best responses and behavioral conjectures

Best response strategies. The theoretical predictions of the Targeting Game, as implemented in our experiment, in terms of best response are the same in the three treatments and they are summarized in Figure 2. Each row in Figure 2 represents one of the 8 blocks, and each column represents one of the 5 periods. There are two rows for each network structure: one where the Opponent targets the node in the center, and one where it targets the periphery. Note that for any of our network structures, the main centrality measures assign the same node as the center.[16] The initial opinion of nodes is not reported in the figure because it does not affect the

[15]Placing the Beauty Contest game before the targeting game may have primed participants' strategic reasoning. However, because no feedback was provided between tasks and participants were informed about their opponent's choices only in the main game, any such effect is likely to have been minimal. We, therefore, acknowledge this potential, but unlikely, source of confounding.

[16]Intermediacy (as defined in online Appendix A.1, equation (7)), closeness centrality (based on proximity to all other nodes), and betweenness centrality (measuring how important a node is in terms of connecting other nodes) give the same unique most central node in each of the four networks. Degree centrality, which takes into account the number of neighbors of the node, displays the same behavior but identifies three central nodes (rather than one) for the line network.

distribution of final opinions (and thus, the strategy profile). The columns in Figure 2 indicate the Opponent's degree of influence λ. In each graph, the Opponent's choice is indicated by a square and the Player's best response (consisting of one or multiple nodes) by a circle.

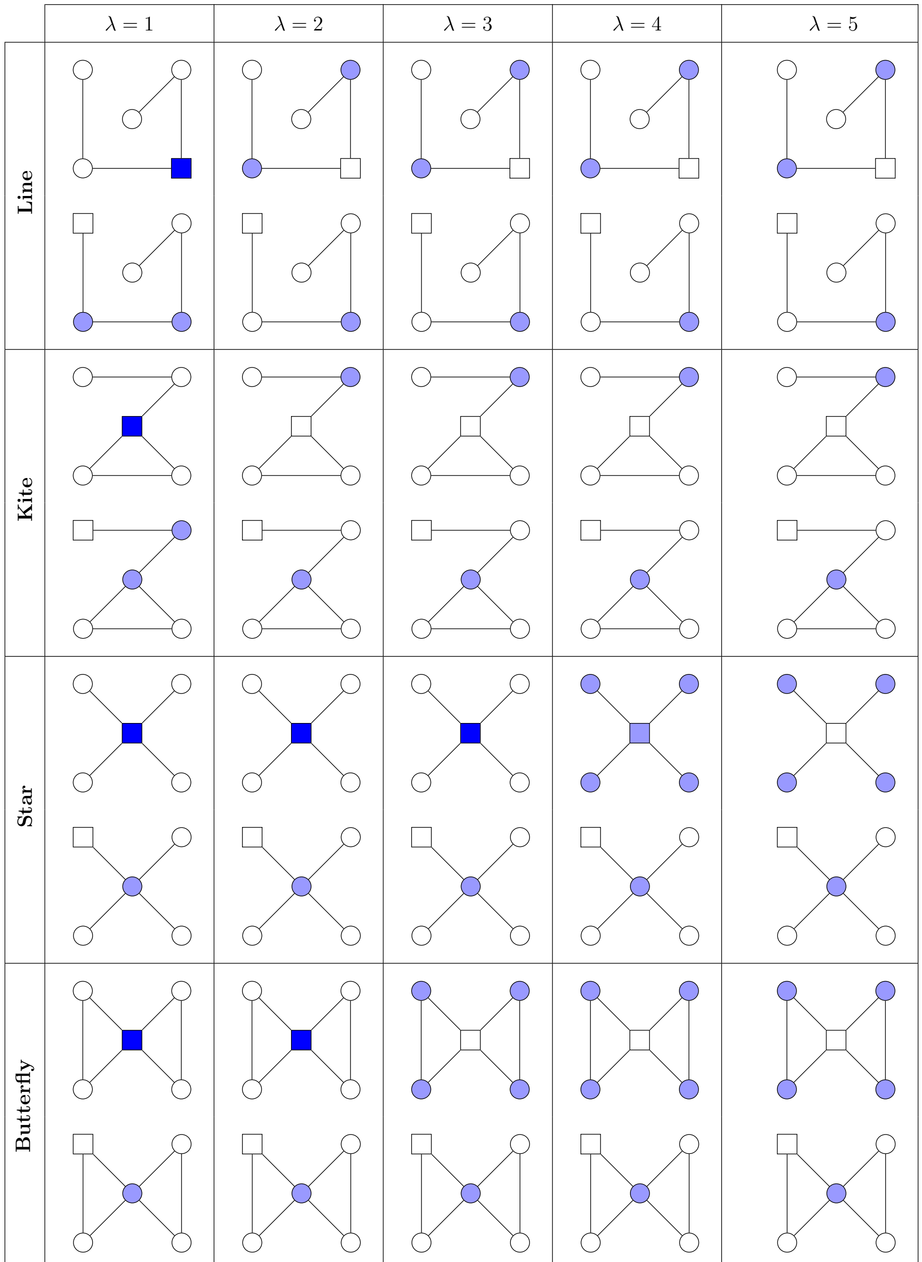
The Player's best response is uniquely identified by the combination of three factors in a scenario defined as the triple (g, s_O, λ), with the network structure g, the observed Opponent's target s_O, and the Opponent's degree of influence λ.[17] As explained in Appendix A, best-response targets are characterized by intermediacy centrality and influenceability (Grabisch et al., 2018), which capture the strategic nature of the two-player game. Intermediacy centrality of a node measures how much a node can interpose itself between the Opponent's target and other nodes, reflecting the extent to which the influence of the Player targeting a given node reaches the network before the Opponent's. Influenceability of a node, on the other hand, given that another node is targeted by the Opponent, is measured by its number of connections and the self-reinforcement of its opinion through the network. The more opinions the node takes into account and integrates with the self-reinforcement process, the less susceptible it is to additional opinion.

The best-response strategy can be summarized as follows: when the Opponent targets a peripheral node, targeting the central node, *i.e.*, the one with higher intermediacy centrality, is always the best response, regardless of the Opponent's influence. Conversely, when the Opponent targets the central node, targeting the center is optimal only when the Opponent's relative influence is low; the Player should switch away from the center when the Opponent's relative influence becomes sufficiently high. When λ is large enough, the Player should target a peripheral, *i.e.*, more influenceable, node, because it is the only one they can effectively influence.

The switch point also depends on the network structure: in the line and kite networks, the center ceases to be the best response for any $\lambda \geq 2$, while for the butterfly and star networks, the transition occurs for $\lambda \geq 3$ and for $\lambda = 5$, respectively. This

[17] Note that if there was only one strategic agent in the Targeting Game, the case would be trivial: in the long run, all nodes' opinions would converge to the strategic agent's opinion because this agent is assumed to be stubborn and does not update, while nodes do.

Figure 2: Best responses in the Targeting Game



Note: The opponent's target is shown as a square, and the player's best response as a blue circle. When only a blue square appears (as in the first line network), it indicates that the player's unique best response coincides with the Opponent's target.

variability in the switching point provides a key source of identification in our design.

Behavioral conjectures. However, individuals may deviate from the best-response strategy and make errors. We consider four possible sources of deviation from best responses.

First, we conjecture that deviations are more likely when decision complexity is greater. In our setting, complexity—defined as the difficulty of identifying the best response—is specific to each scenario, and thus is defined at the level of a triple (g, s_O, λ). That is, no network configuration, opponent choice, or value of λ is inherently difficult per se; rather, complexity arises from the interaction of these three dimensions within a given scenario, which may make subjects less likely to identify the best response. However, some features may be a source of complexity. This is in particular the case for the distance to the switching point. In the game, when the Opponent targets Periphery, the best reply is always Center, but when the Opponent targets Center, the switching point is defined as the smallest λ at which the Center ceases to be the best response. As such, this variable captures an additional dimension of strategic complexity. Indeed, in some configurations, the node that corresponds to the best response changes as λ varies, whereas for others it does not. Also, choices may be more or less difficult when λ is just above the cutoff than when it is substantially larger. The number of best responses in the network could also be a source of complexity because if multiple best responses introduce a slack (choosing a best response node is more likely), the absence of a unique best response may reduce clarity if participants search for a unique best node.[18]

Our empirical analysis proceeds in two steps. First, we identify and test several features that proxy for complexity, including network configuration, the number of best replies, and the distance from the switching point. Second, we include scenario fixed effects, which absorb complexity at its most granular level.

Second, we conjecture that deviations from best responses are more prevalent among participants with lower cognitive abilities or less experience in the game. In

[18]Network size could also be a source of complexity, but we cannot test this conjecture because network size does not vary in our study.

particular, participants may make more errors in the earlier blocks where learning takes place than in the later ones.

Third, we conjecture that behavioral biases and heuristics can lead to deviations from best responses. Specifically, some players may target focal points (*e.g.*, the center of a network) when selecting a node. Similarly, some players may avoid selecting the Opponent's target, even when this is not the optimal choice. Another possible cause of sub-optimality is targeting like-minded nodes in heterogeneous networks, although initial opinions are irrelevant in terms of best response.

Finally, we test whether behavior is invariant to the theoretically irrelevant order in which the Opponent's degree of influence changes and to the visual presentation of the network. Behavior should be independent of sequencing or cosmetic display features. Any systematic deviation from invariance would suggest behavioral forces such as salience, anchoring, or path dependence. If anchoring plays a role, early exposure to strong influence asymmetry in the Decreasing Influence treatment may bias behavior more than in the Increasing Influence treatment, where players start in a symmetric position. If visual presentation matters, a plausible directional mechanism is salience: participants may disproportionately select nodes displayed at the center of the screen, making more mistakes in the layout treatment when targeting the central node is optimal.

3.3 Procedures

A total of 19 sessions were run at GATE-Lab in Lyon, France, with 423 participants: 11 sessions with 220 participants in the main Increasing Influence treatment, 4 sessions with 99 participants in the Decreasing Influence treatment, and 4 sessions with 104 participants in the Layout treatment.[19] Participants were primarily students from local engineering, business, and medical schools, recruited via HRoot (Bock

[19] The number of observations in the main treatment (220) was based on an ex-ante power calculation, aiming for a power of 0.95 to detect a small effect size (Cohen's d = 0.25), using two-tailed Wilcoxon signed-rank tests and assuming a Type-I error rate of 0.05. For the control treatments that we conducted after analyzing the results of the main experiment, we targeted twice fewer observations to detect a medium-size (d=0.5) effect with a power of 95% in two-tailed Mann-Whitney tests comparing treatments. We slightly deviated from this number due to no show-up.

et al., 2014).[20] The experiment was programmed using Java.

Upon arrival at the laboratory, participants were randomly assigned to a computer terminal. They received written instructions for each of the first two parts (the Beauty Contest game and the Targeting Game, respectively) after they completed the previous step. Instructions were read aloud (see Appendix B.1). Those for the Targeting Game included several illustrations to help understand the weights of the nodes, the Opponent's degree of influence, and the process of opinion revision.

Before starting part 2, we checked the understanding of the instructions using a quiz in which participants had to calculate the weight of the opinion of each node on each other node in various configurations, and the weights of the player's and opponent's opinions on various nodes in given examples (see questions in Appendix B.1). In case of a wrong answer, participants could enter new responses without any time limitation (the quiz contained 17 questions in total, with minimum number of mistakes = 0, max = 59, mean =10.9, S.D. = 10.2). We checked individually each participant's understanding, answered their questions, and reexplained the instructions when needed until they were able to complete the quiz correctly. We did not let a subject start the experiment before making sure that they were able to answer the questions correctly. Then, as already mentioned, to familiarize them with the environment, participants had 6 minutes to practice the interface and observe the opinion updating process within the network. During this practice, they could observe a tree network, different from the network structures used in the experiment, with both homogeneous and heterogeneous initial opinions. They could change their own targeting choice, the Opponent's choice, and its degree of influence. Payoffs in this practice were hypothetical but players were informed about their hypothetical payoff. During the game, participants' decisions were not time-constrained.

The instructions for the Raven matrices were displayed on the screen. After part 3, we asked socio-demographic questions and queries about their social media usage, risk attitudes, and social preferences. Finally, we asked "What is the perceived degree of difficulty of the decisions in Part 2 on a scale of 0 (very simple) to 10 (very

[20]See Table D1 in online Appendix D for summary statistics of the socio-demographic characteristics of the participants. They do not significantly differ across treatments, except for age.

complex)?" (mean =4.5, S.D. = 2.5) and "Of your 40 decisions in Part 2, how many did you make at random?" (mean =7.5, S.D. = 8.6).

At the end of the session, the program randomly selected three periods in different blocks of part 2 and added the earnings in these three periods. Participants earned an average of €22.7 (S.D.= 2.0) for their decisions, including a €5 show-up fee. Each session lasted about 1.5 hours. Participants were paid privately in a separate room.

4 Results

We begin by reporting the frequency of node choices and best responses across conditions. Next, we explore the determinants of choosing specific nodes in the network to identify behavioral regularities. Finally, we examine efficiency.

4.1 Targeting choices and best responses

Descriptive statistics. Figure C2 in Appendix C adds to Figure 2 by showing the relative frequency with which each node was selected under each condition of the Increasing Influence treatment.[21] It shows that a majority of players targeted a node that was not chosen by their opponent (represented by a rectangle), except when the opponent targeted the central node and the influence degrees were relatively balanced. In many cases, the modal choice is the node in the center, which frequently corresponds to a best response in networks like stars and butterflies.

Moving from relative frequencies of node choices to best responses, Table 1 summarizes the percentage of targeting choices that were a best response in each configuration of the Increasing Influence treatment. Participants chose the best response (or one of them, when multiple) 70% of the time. There are no significant differences when compared to the 69% frequency in the Decreasing Influence treatment (two-tailed Mann-Whitney rank-sum test, MW hereafter; $p = 0.204$) and the 68%

[21]Figures C3 and C4 provide similar information in the two control treatments.

frequency in the Layout treatment (MW, $p = 0.279$).[22,23]

Table 1: Relative frequency of best responses in the Increasing Influence treatment

All networks	0.7				
Opponent's target	Center	Periphery			
	0.65	0.76			
Best response	Center	Periphery			
in the network	0.75	0.59			
Opponent's degree	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$	$\lambda = 5$
of influence	0.73	0.56	0.66	0.82	0.75
Network structure	Line	Kite	Star	Butterfly	
	0.67	0.64	0.77	0.75	
# Best responses	Unique	Multiple			
in the network	0.71	0.69			
Nodes' initial	Homogen.	Heterog.			
opinion	0.72	0.69			

Notes: The table reports the relative frequencies of best responses. *P*-values from Wilcoxon signed-rank tests for each possible pairwise comparison can be found in the Table D3 in Appendix D.

Table 1 shows interesting differences in the best response rate across conditions. This rate is higher when the opponent targets a peripheral node rather than a node at the center (76% *vs.* 65%; Wilcoxon signed-rank test, W hereafter, $p < 0.001$).[24] It is also significantly higher when the center of the network is the best response (75% *vs.* 59%; W test, $p < 0.001$). The best response rate varies by network structure. Participants were more likely to best respond in the star and butterfly networks (77 and 75%, respectively, compared to 64% and 66% in the kite and line networks), that is, in settings where the network structure has a focal point.[25,26]

[22]See Table D2 in the Appendix D for between-treatment comparisons. The main significant differences in the best response rate between treatments are found for the λ parameter (at the 1% level when $\lambda > 1$). This is likely driven by the different points in time at which participants need to switch targets to best respond.

[23]Throughout this section, we used Mann-Whitney rank-sum tests for between-subject comparisons (independent samples) and Wilcoxon signed-rank tests for within-subject comparisons (matched samples). All tests are two-tailed.

[24]There are no significant differences across treatments (see Table D2, MW tests, $p > 0.22$).

[25]Table D3 in Appendix D indicates that all pairwise comparisons are highly significant (W test, $p < 0.001$), except the differences between line and kite (W test, $p = 0.088$), and between star and butterfly (W test, $p = 0.057$).

[26]The Layout and Decreasing Influence treatments differ significantly from the Increasing Influ-

The rate also varies depending on the opponent's degree of influence, but non linearly so;[27] this should be interpreted with care since λ varies with the period number and the treatment. Finally, it slightly differs according to whether the nodes have homogeneous or heterogeneous initial opinions (72% *vs.* 69 %; W test, $p = 0.010$).[28] In contrast, the rate is not significantly higher in networks with multiple best responses compared to those with a unique best response (69% *vs.* 71%, respectively; W test, $p = 0.066$).

Patterns of observed choices. Could the high percentage of best responses be due not to optimal reasoning but to chance, notably because some networks had more than one best response node? As Figure 2 shows, only a minority of configurations admit multiple best responses (11 out of 40 networks); overall, there are 62 best-response nodes out of 200 (instead of 40 if only unique best responses were admitted). Thus, the unconditional probability of targeting a best-response node by choosing at random is $62/200 = 31\%$, while participants best-responded in 69% of instances across all treatments. This suggests that our results are unlikely to be driven by random targeting at the level of the observed choice. Because the best response is defined at the level of Network $\times$ Opponent's target $\times$ λ, we also compared observed choices to random choices at this same level. We conducted 40 Wilcoxon rank-sum tests contrasting, for each scenario (Network $\times$ Opponent's target $\times$ λ), the binary indicator of whether the participant chose the best response with its analogue under a purely random scenario in which each node is selected with a 20% probability. All tests reject equality except for one scenario with four best responses.[29]

We also rule out the possibility that the high best response rate is due to a general attraction to the central node (which is more frequently a best response, especially in the butterfly and star networks). Although 75% of participants best

ence treatment only in line networks (MW test, $p = 0.022$ and 0.002 respectively) (Table D2).

[27] Table D3 in Appendix D indicates that all pairwise comparisons between λs are highly significant (W test, $p < 0.001$), except the difference between λ=1 and λ=5 (W test, $p = 0.528$).

[28] Table D2 shows no significant differences across treatments (MW tests, $p > 0.20$).

[29] In 38 out of 40 scenarios we get p-value< 0.001. In the Butterfly network when the Opponent targets center, for $\lambda = 4$ and $\lambda = 5$ we get p-value $= 0.576$ and 0.038, respectively.

responded when the center was the optimal target, 59% also best responded when the best response lay in the periphery (Table 1). Target choices display substantial variation: none of the 423 participants targeted the center in every block, and only 22 did so in more than five blocks. On average, players targeted 2.1 distinct nodes per block, and targeted the center in all five periods in only 27% of blocks. Behavior responds to the Opponent's choice, as predicted by the model (see Figure 2). When the Opponent targeted the center, players visited 2.4 nodes on average and chose the center exclusively in only 12% of blocks. However, even when the Opponent targeted a peripheral node (making the center the best response), players still visited 1.8 nodes on average per block, with only 43% of blocks showing persistent center targeting.

As seen in Table 1, most of the choices can be rationalized as best responses (for an analysis of the instances where players did not best respond, see Figure 5 in subsection 4.3). Nearly 70% of suboptimal choices occurred when participants targeted a different node than their Opponent, especially when the Opponent targeted the center, even when it would have been optimal to do otherwise. Overall, these elements rule out that many players played randomly or used the heuristics of always targeting the center to cope with the game's complexity. The relatively high frequency of best response above chance, even when it did not correspond to the center, rather suggests that most participants understood the game well.

Regression analysis. We explore the determinants of whether participants played the best-response strategy, controlling for the characteristics of players and games. Pooling the data from the three treatments[30] gives a sample of 16,920 unique decision sets (423 participants X 8 blocks X 5 periods) where participants had to target one

[30] We pool these data because best responses are identical across treatments: for any given network configuration, opponent target, and value of λ, the best-response mapping is the same. Empirically, Table D2 shows no significant differences in best-response rates across treatments when aggregated over λ, networks, and opponent targets; the deviations observed at specific λ values simply reflect that participants encountered these values at different periods in the block. Once we control for block number or include fixed effects for each (Network X Opponent's target X λ) cell, these differences disappear. Because identification comes from variation in these elements (and we include fixed effects for all of them), pooling preserves internal validity, with any residual treatment-order or layout differences absorbed by treatment dummies.

node out of five. We estimated a linear probability model specified as follows:

$$BestResponse_{i,d} = \alpha DIT_d + \beta LayoutT_d + \delta Heterogenous_d + \gamma X_i + \kappa Block_d + \epsilon_{i,d} \quad (1)$$

where $BestResponse_{i,d}$ is the dependent variable that is equal to 1 if the choice of player i for decision set d is a best response, and 0 otherwise. Table 2 reports the estimates of this model.[31]

All specifications include dummies for the Decreasing Influence and Layout treatments, with the Increasing Influence treatment as the reference category. We also include a dummy indicating whether initial opinions were heterogeneous (the reference category is the homogeneous case in which all initial opinions equal 0.5). X_i represents a vector of individual characteristics. To capture abilities relevant for mathematical, cognitive, and strategic reasoning abilities, we control for whether the participant attends the Central Engineering School (a highly selective, math-intensive program), the Raven score (0-6), and the number chosen in the Beauty Contest game. We further include three characteristics: risk attitude, social media use (a dummy equal to 1 for moderate use, below one hour per day, as a proxy for networking experience), and pro-sociality (rescaled so that one unit corresponds to a 10-point donation). We also control for block number (1-8) to capture learning effects. The error term $\epsilon_{i,d}$ is clustered at the participant level.

We estimated three specifications that vary the fixed effects included. The specification of column (1) reflects Eq. (1). In column (2), we added controls capturing specific aspects of the scenario complexity: dummies for each network structure (with the line network as the omitted category), the Opponent's choice (equal to 1 if the Opponent targeted the periphery, 0 otherwise), the number of best responses (1-5), and the distance from the switch point.[32] We included fixed effects for periods (1-5). In column (3), we replaced the controls of column (2) with a set of fixed effects repre-

[31] Estimating a Probit model with similar independent variables as the linear probability model provides qualitatively similar results. They are omitted but can be provided upon request.

[32] We define the switch point as the smallest λ such that Center ceases to be the best response when the Opponent targets Center. As an example, in the Butterfly network, the distance from the switch point is $-2, -1, 0, +1, +1$ for $\lambda = 1, ..., 5$, respectively, when the Opponent targets Center. Conversely, the switch point is always zero when the Opponent targets Periphery.

senting the combination of network configuration, the Opponent's choice (center *vs.* periphery), and λ (40 effects).[33] Since these three factors uniquely identify the best response and the complexity-relevant scenario, this is our preferred specification.

Suggesting a role of complexity, results from column (2) in Table 2 indicate that identifying the best response was more likely when the network structure had a clear focal point and several best responses (star and butterfly networks), as already visible in Table 1. Pinning down the best response tends to be easier in the presence of multiple best responses (only marginally significantly so), for a value of λ above the switch point, and when the Opponent targeted the periphery.

Table 2 also provides evidence that mathematical, cognitive, and strategic reasoning abilities are significantly associated with the likelihood of playing a best response. These abilities, proxied by studying at the Central engineering School, solving a greater number of Raven matrices, and reporting a lower number in the Beauty Contest game, increased the probability of best-response play, probably because they facilitate strategic reasoning.[34] Results regarding the block number suggest that some learning occurred during the game progression.[35]

Table 2 also shows that a best response was significantly less likely to be chosen when the initial opinions of nodes were heterogeneous, a result that is inconsistent with the theory and that we explore in a later section. Finally, it indicates no significant treatment differences, consistent with the theory.

All the above results hold when controlling for all fixed effects at the level of *Network X Opponent's target X* λ in column (3), which is the most conservative specification. This analysis supports our first result:

Result 1: *A majority of participants chose the best response in the Targeting Game, with this tendency being even more pronounced in less complex scenarios and among those with stronger mathematical, cognitive, and strategic reasoning abilities.*

[33]As in some sessions λ decreased over periods, the effect of periods can be identified separately from λ.

[34]Prosociality was also significantly associated with a lower likelihood of best responding (perhaps due to less strategic thinking). In contrast, risk attitude and social media use were not significant.

[35]We also found a slightly negative effect of extended decision time on the probability of a best-response. The mean decision time was 7.4 sec. (maximum = 5.7 min.; median = 3.5 sec.)

Table 2: Determinants of best responses

Dependent variable: Choosing the best response	(1)	(2)	(3)
Increasing Influence treatment	*Ref.*	*Ref.*	*Ref.*
Decreasing Influence treatment	-0.028*	-0.028*	-0.028*
	(0.016)	(0.016)	(0.016)
Layout Treatment	-0.029	-0.029	-0.029
	(0.018)	(0.019)	(0.019)
Heterogeneous opinion	-0.027***	-0.027***	-0.022***
	(0.008)	(0.008)	(0.007)
Engineering school	0.048***	0.048***	0.048***
	(0.014)	(0.014)	(0.014)
Raven score	0.022***	0.022***	0.022***
	(0.005)	(0.005)	(0.005)
Number in the Beauty Contest	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)
Social media $<$ 1H/day	0.010	0.010	0.010
	(0.016)	(0.016)	(0.016)
Risk attitude	-0.001	-0.001	-0.001
	(0.004)	(0.004)	(0.004)
Prosociality (1 unit=10)	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)
Block number	0.017***	0.018***	0.018***
	(0.002)	(0.002)	(0.002)
Network: Kite	-	0.007	-
		(0.011)	
Network: Star	-	0.159***	-
		(0.012)	
Network: Butterfly	-	0.105***	-
		(0.012)	
Opponent choice: Periphery	-	0.118***	-
		(0.011)	
Number of best responses	-	0.011*	
		(0.006)	
Distance from switch point	-	0.018***	
		(0.004)	
Constant	0.635***	0.478***	0.470***
	(0.037)	(0.042)	(0.044)
F.E. period	NO	YES	YES
F.E. Network X Opponent's target X λ	NO	NO	YES
Observations	16,920	16,920	16,920
R-squared	0.025	0.073	0.144

Notes: This table reports regressions from a linear probability model. Robust standard errors, clustered by participant, are in parentheses. FE for fixed effects. *** p<0.01, ** p<0.05, * p<0.1.

4.2 Choosing a target

We now investigate behavioral tendencies underlying the selection of a specific target beyond the best response.

Behavioral regularities. To uncover systematic tendencies, we analyzed the factors influencing which node was targeted. With $16,920$ decision sets (423 participants X 8 blocks X 5 periods), we observe $16,920 * 5$ nodes $= 84,600$ choices. Our binary dependent variable, $choice_{n,d}$, takes value 1 if node $n = 1, ..., 5$ is targeted in a given decision set d. Since only one node is selected in each decision set, we observe one non-zero outcome out of five. Thus, we estimated a conditional Logit model that takes into account the dependence between the different options, specified as follows:

$$P(\text{choice}_{n,d} = 1) = \frac{\exp(\beta \cdot X_{n,d})}{\sum_{j=1}^{5} \exp(\beta \cdot X_{j,d})} \tag{2}$$

where $P(\text{choice}_{n,d} = 1)$ is the probability that node n is chosen in decision set d. $X_{n,d}$ represents the vector of characteristics for node n in decision set d, and β is the vector of coefficients to be estimated. The denominator sums over the five nodes in the decision set d, ensuring the probabilities sum up to 1. This model allows for dependency across options within a given decision set of five alternatives, which lets us capture the relative utility of each node based on its attributes. Standard errors are clustered at the individual level.

Table 3 reports the estimates from four specifications of this regression model. The corresponding average marginal effects are reported in Appendix Table D4. The first three specifications include three dummy variables equal to one if the node is the best response, if the node is targeted by the Opponent, and if the node is at the center of the network.[36] Columns (2) and (3) also include the node's initial opinion (0.25 and 0.75, with 0.5 as the omitted category). In column (3), the sample is restricted to blocks with heterogeneous initial opinions of nodes. The specification in column (4) excludes cases where the best response is the center to test whether

[36]Within a given decision set, the factors jointly determining the best response (network, Opponent's target, and λ) stay the same. This is why the fixed effects at this level are not included.

individuals were still more likely to choose the best response.[37]

Table 3: Conditional Logit estimates of targeting a given node

Dependent variable: Choice of a node	(1) All blocks	(2) All blocks	(3) Heterogeneous blocks	(4) Best response is not center
Best response	1.159***	1.173***	1.166***	0.806***
	(0.038)	(0.038)	(0.044)	(0.053)
Node targeted by Opponent	-0.286***	-0.304***	-0.463***	-
	(0.039)	(0.040)	(0.051)	
Node in the center	1.196***	1.229***	1.153***	-
	(0.040)	(0.040)	(0.050)	
Node's initial opinion: 0.25	-	-0.250***	-0.238***	-0.225***
		(0.056)	(0.054)	(0.078)
Node's initial opinion: 0.75	-	0.488***	0.478***	0.462***
		(0.052)	(0.051)	(0.069)
Observations	84,600	84,600	42,300	25,380
Pseudo-R-squared	0.276	0.286	0.269	0.050
Log-pseudo likelihood	-19705.9	-19440.8	-9955.8	-7764.9

Notes: This table reports the estimated coefficients from conditional Logit estimates by decision set (participant X block X period). Robust standard errors, clustered at the individual level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

All regressions in Table 3 show that a node is significantly more likely to be selected when it corresponds to the best response, including when the best response is not the center of the network (see column (4)), showing the understanding of the game. Nevertheless, even after controlling for the best response, a residual tendency remained to target the center rather than the periphery and to avoid targeting a node selected by the Opponent. This analysis yields our second result:

Result 2: *Even after accounting for best-response incentives, participants showed a residual tendency to target the central node and to avoid the node targeted by the Opponent.*

[37] In our setup, the center is not the best response only in a subset of games where the Opponent targets the central node. Thus, the variables indicating whether the Opponent targets this node or the node is in the center are excluded because of collinearity with the best response.

Heterogeneity analysis. As a robustness check, we re-estimated the specification from column (2) of Table 3, splitting the sample along various individual and game characteristics.[38] Appendix Table D5 separates blocks with homogeneous *vs.* heterogeneous opinions (columns (1)–(2)); the Increasing, Decreasing, and Layout treatments (columns (3)–(5)); and earlier *vs.* later blocks (columns (6)–(7)). Table D6 splits participants by engineering-school status (columns (1)–(2)); below-*vs.* above-average Raven scores (columns (3)–(4)); below- *vs.* above-average self-reported number of random choices (columns (5)–(6)); and below- *vs.* above-average number of mistakes in the understanding quiz (columns (7)–(8)).[39] Across all splits, the results closely mirror those in Table 3, showing that our main findings are not driven by specific subgroups or game conditions.

Affinity and opposition biases. Table 3 shows that nodes' initial opinions affected targeting, although theoretically irrelevant. Participants were significantly more likely to target nodes closer to their own opinion (0.75), and less likely to target those closer to the Opponent's (0.25), compared to neutral nodes (0.50). Regressions (8) to (11) in Table D5 show similar patterns for engineers and non-engineers, and for participants with both high and low Raven scores (the effect is weaker among the latter), indicating that this behavior is not solely driven by lower cognitive ability.

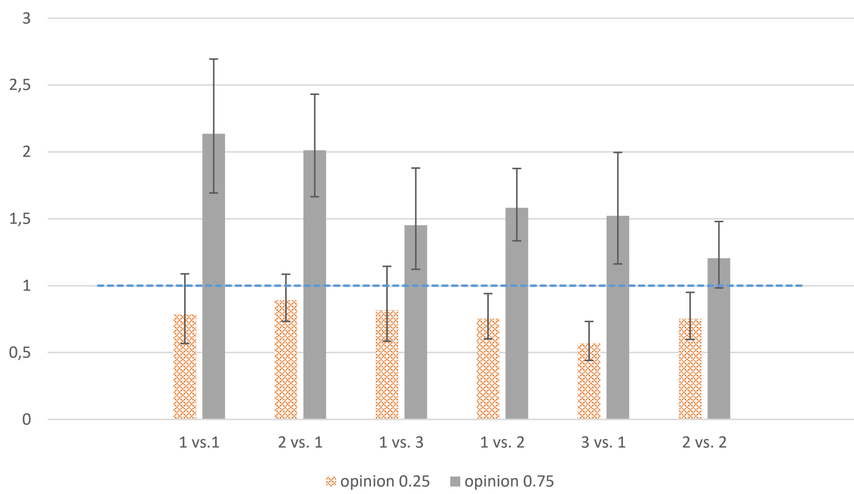
The specifications in Table 3, however, do not account for how the distribution of initial opinions may shape targeting. To recall, in each heterogeneous block, the three initial opinion values (0.25, 0.5, and 0.75) were randomly assigned so that each appeared at least once, generating six possible distributions, denoted by the number of nodes holding 0.25 and 0.75 opinions (*e.g.*, '1 *vs.* 3' means one node with 0.25, three with 0.75, and one with 0.5). Table D7 in the Appendix estimates how the effects of opinions 0.25 and 0.75 vary across these distributions. Column (1) reports conditional logit coefficients for the full sample; column (2) restricts to heterogeneous blocks; and column (3) reports the odds ratios from the regression of column (2).

[38] Estimating a linear probability model with similar specifications delivers qualitatively similar results. They can be provided upon request.

[39] For the definition and descriptive statistics of these variables see Section 3.3.

Figure 3 plots these odds ratios. The omitted category (opinion 0.5) is depicted by a horizontal line at 1. Orange (dotted) bars represent the odds ratios for nodes with opinion 0.25, and dark-gray bars to those with opinion 0.75.

Figure 3: Impact of the distribution of initial opinions in the network on the selected node



Notes: This figure reports the odds ratios from the conditional logit estimates in Appendix Table D7. Reading: *e.g.*, "2 *vs.* 1" reads as two nodes with initial opinion 0.25 and one with opinion 0.75. To recall, the player's opinion is always 1. Error bars show the 95% confidence intervals: when a bar crosses the horizontal line, the implicit targeting probability is not statistically distinguishable at the 95% confidence level from that associated with nodes with opinion 0.5.

Overall, the dark-gray bars indicate a strong tendency to target nodes closer to one's own opinion ('affinity bias'), whereas the light bars reveal a moderate avoidance of more distant opinions ('opposition bias'). The strength of these biases depends on the opinion distribution. Biases are weakest in balanced configurations (2 *vs.* 2, far right of the figure), where affinity bias disappears and only a mild opposition bias remains. Conversely, they are strongest when the closer opinion is under-represented (3 *vs.* 1), producing both affinity bias and the largest opposition bias. When only one node shares the closer opinion but the distant opinion does not form a majority (1 *vs.* 1; 2 *vs.* 1; far left of the figure), affinity bias is pronounced while opposition bias is not significant.[40] This analysis leads to our last result:

Result 3: *There is a strong preference for targeting nodes with initial opinions closer to one's own and a moderate aversion to targeting nodes with opinions closer to that of the Opponent. These affinity and opposition biases are lowest in a balanced setting where both opinions are represented by multiple nodes, and they are highest when the closer opinion looks vulnerable because it is represented by one node only.*

4.3 Efficiency

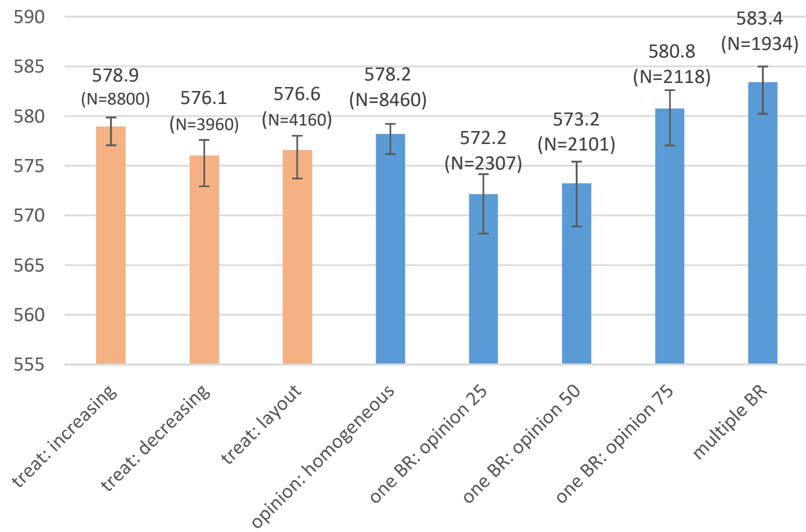
How costly are the behavioral biases in terms of efficiency, as measured by the participants' earnings? Participants were paid proportionally to the maximum attainable payoff, with best responders always earning 600 ECU, which serves as our natural efficiency benchmark. Figure 4 shows average payoffs across all games and players by condition, with 95% confidence intervals, illustrating how far earnings fall short of the hypothetical 600 ECU. Consistent with Table 1, payoffs are slightly higher in the Increasing Influence treatment than in the Decreasing Influence and Layout treatments, although differences are not significant at standard levels (MW tests).[41]

Since initial opinions were randomly distributed and did not influence best responses, any differences in average payoffs directly reflect behavioral biases in targeting. Average payoffs in the homogeneous condition closely match those in the Increasing Influence treatment. In the heterogeneous condition, we further split games by the initial opinion of the best-response node(s). Bars labeled 'one BR: opinion 25', 'one BR: opinion 50', and 'one BR: opinion 75' correspond to cases in which the unique best-response node had opinion 0.25, 0.50, or 0.75, respectively; the last bar pools cases with multiple best responses, which mechanically yield higher payoffs because the latitude for errors is reduced.

[40] Table D8 in the Appendix reveals that if they exhibit the same type of heuristics, the low-Raven participants are systematically more likely to target a node whose initial opinion is closer, regardless of the network configuration in terms of initial opinions, while high-Raven participants react more to some configurations.

[41] In all treatments, realized payoffs exceed the expected 533 ECU under random node selection. Individual earnings also vary with cognitive ability: splitting the sample at the average Raven score, high-Raven participants earn, on average, 4.9, 8.5, and 3 ECU more than low-Raven participants in the Increasing, Decreasing, and Layout treatments, respectively.

Figure 4: Average payoffs across treatments and conditions



Notes: This figure reports average payoffs (ECU) by condition, with 95% confidence intervals. The unit of observation is the targeting choice (N=16,920). The first three bars correspond to the Increasing, Decreasing Influence, and Layout treatments, respectively. The 4th bar shows the homogeneous condition; the next three bars correspond to the heterogeneous condition with a unique best response ("BR") of opinion 0.25, 0.5, and 0.75, respectively. The far-right bar is for the heterogeneous condition with multiple best responses.

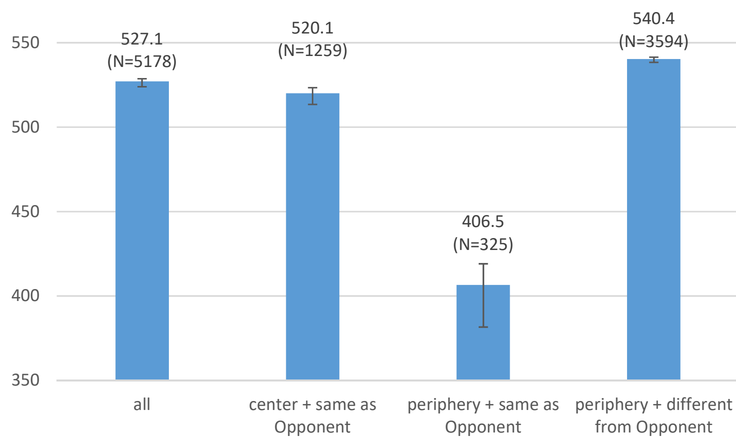
In the heterogeneous condition, average payoffs were lowest when the best response was a node with the most distant opinion (mean= 572.2) and highest when it was the closest (mean= 580.8). This difference is significant (MW test, $p < 0.001$), and both means are significantly below the 600 ECU benchmark (t-test, $p < 0.001$). Thus, when the best-response node had opinion 0.75, participants earned on average 8 ECU points more than when the best response had opinion 0.25, providing a measure of the efficiency cost of affinity and opposition biases.

Figure 5 focuses on the 5178 decisions that deviated from the best response. We classify deviations into three types of 'wrong' choices: targeting the center like the Opponent (1259 choices, 24.3%), targeting a peripheral node like the Opponent (325 choices, 6.3%), and targeting a peripheral node when the Opponent targeted the center (3594 choices, 69.4%).[42] Most efficiency losses arise when participants

[42] When the Opponent targets the periphery, targeting the center is always optimal. Thus, "Cen-

targeted a node different from their Opponent, particularly when they targeted the periphery while the Opponent targeted the center. Targeting the center when it is not the best response yielded only 520.1 on average, while targeting the periphery when the Opponent targeted the center yielded 540.4 on average, both significantly below 600 ECU (t-test, $p < 0.001$). Although infrequent (N=325), following the Opponent in targeting a peripheral node produced the lowest earnings (406.5).

Figure 5: Average payoffs when the selected target is not the best response



Notes: This figure shows average payoffs when participants did not choose the best response, with 95% confidence intervals. The unit of observation is the individual targeting choice. The 2nd and 3rd bars correspond to incorrect targeting of the center or periphery, respectively, that mirrors the Opponent's choice. The far-right bar captures incorrect peripheral targeting when the Opponent targeted the center.

5 Discussion and conclusion

We experimentally examined how individuals select a target node to maximize their influence on average public opinion in a network when their opponent's target is known. We conjectured four potential sources of deviations from best responses: complexity, limited cognitive and strategic abilities, heuristics, and sensitivity to

ter + different from Opponent" is not a category in this figure.

irrelevant design features. Participants best responded to their opponent's targeting choices approximately 70% of the time. This likelihood decreased with complexity and lower cognitive and strategic abilities, while changes in game layout or ordering had no significant effect. Despite this relatively high rate of optimal play, systematic behavioral deviations emerged. Conditional on not best responding, we found a residual tendency to target the center of the network rather than its periphery and avoid nodes already selected by the opponent, suggesting a differentiation strategy.

One intriguing finding is the evidence of affinity and opposition biases. Players were more likely to target nodes whose initial opinion was closer to their own and to avoid nodes closer to their opponent, although initial opinions were payoff-irrelevant. These biases were strongest in unbalanced settings.

These biases may stem from two main mechanisms. First, they may be consistent with homophily (*e.g.*, McPherson et al., 2001; Currarini et al., 2009; Benhabib et al., 2010; Golub and Jackson, 2012; Currarini and Mengel, 2016; Acemoglu et al., 2021), or identity-based mechanisms if similar-minded nodes were perceived as in-group (*e.g.*, Akerlof and Kranton, 2000; Chen and Chen, 2011; Chen et al., 2014). This explanation would require that participants treated robots like humans. Second, participants may have mistakenly believed that more similar-minded nodes are easier to influence. These tendencies appear unrelated to comprehension: coefficients remain significant for both high- and low-Raven participants (models (10)-(11) in Appendix Table D5). Although biases are weaker for higher Raven-score players, they persist despite a correlation of -0.244 between comprehension errors in the questionnaire and the Raven score. Evidence of learning is limited, as avoidance of dissimilar nodes does not change over time (models (6)-(7) in Table D5). These patterns do not require subjects to fully grasp the theoretical irrelevance of initial opinions. Our interpretation is descriptive: whether driven by correct or mistaken beliefs, the systematic preference for like-minded nodes reflects a robust behavioral heuristic rather than random confusion.

These results and the fact that network composition shapes the strength of affinity and opposition biases have efficiency implications: if influencers value opinion proximity in contexts such as social media or coalition formation, more balanced networks with diverse initial opinions may mitigate targeting distortions.

Our study offers a first step toward understanding targeting competition, an important but largely overlooked topic, and we acknowledge its limitations. Targeting decisions were implemented sequentially against an automated opponent, removing the strategic uncertainty central to the simultaneous-move game of the theoretical model. Therefore, our results capture failures to best respond in a deterministic environment but cannot address other mechanisms that could generate deviations from the equilibrium in a full strategic setting, such as ambiguity or risk aversion, biased belief formation, and mistakes in anticipating an opponent's choice. Still, isolating best-response behavior in a transparent environment provides a useful benchmark, revealing which deviations emerge even in the absence of uncertainty. Future work could introduce first a human opponent in a sequential game, and then a simultaneous version, to study how additional strategic forces shape behavior. This would generate mixed-strategy equilibria and require deeper reasoning, probably amplifying behavioral biases.

Our networks were simple and perfectly observable. The fact that we observed systematic behavioral deviations even in this transparent setting suggests that such mistakes may be even stronger in environments with larger, more complex, and only partially known networks. Future work could vary these dimensions. Moreover, exploring other updating mechanisms than the DeGroot rule, weighting neighbors differently, incorporating distance from strategic agents, or allowing for memory failures would broaden the scope of the analysis. Finally, an important avenue for further research is disentangling homophily from the belief that like-minded nodes are easier to influence in networks.

References

Acemoglu, D., A. Ozdaglar, and J. Siderius (2021). Misinformation: Strategic sharing, homophily, and endogenous echo chamber. *NBER Working Paper 28884*.

Akbarpour, M., S. Malladi, and A. Saberi (2025, November). Just a few seeds more:

Value of network information for diffusion. *American Economic Review 115*(11), 3713–48.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Ballester, C., A. Calvo-Armengol, and Y. Zenou (2006). Who's who in networks. Wanted: The key player. *Econometrica 74(5)*, 1403–1417.

Ballester, C., A. Calvo-Armengol, and Y. Zenou (2010). Delinquent networks. *Journal of the European Economic Association 8*, 34–61.

Banerjee, A., A. Chandrasekhar, E. Duflo, and M. Jackson (2013). The diffusion of microfinance. *Science 341, DOI: 10.1126/science.1236498*(6144).

Banerji, A. and B. Dutta (2009). Local network externalities and market segmentation. *International Journal of Industrial Organization 27*, 605–614.

Battiston, S., M. Puliga, R. Kaushik, P. Tasca, and G. Caldarelli (2012). Debtrank: Too central to fail? Financial networks, the FED and systemic risk. *Nature Scientific Reports 2*, 541.

Beaman, L., A. BenYishay, J. Magruder, and A. M. Mobarak (2021). Can network theory-based targeting increase technology adoption? *American Economic Review 111*(6), 1918–1943.

Bechler, C. J., Z. L. Tomala, and D. D. Rucker (2020). Choosing persuasion targets: How expectations of qualitative change increase advocacy intentions. *Journal of Experimental Social Psychology 86*, 103911.

Belhaj, M. and F. Deroïan (2019). Group targeting under networked synergies. *Games and Economic Behavior 118*, 29–46.

Belhaj, M., F. Deroïan, and S. Safi (2023). Targeting in networks under costly agreements. *Games and Economic Behavior 140*, 154–172.

Benhabib, J., A. Bisin, and M. O. Jackson (2010). *The Handbook of Social Economics*, Volume 1A and 1B. Amsterdam: North Holland.

Bimpikis, K., A. Ozdaglar, and E. Yildiz (2016). Competitive targeted advertising over networks. *Operations Research 64(3)*, 705–720.

Bloch, F. (2016). Targeting and pricing in social networks. In Y. Bramoullé, A. Galeotti, and B. W. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, pp. 776–791. Oxford University Press.

Bloch, F. and S. Shabayek (2023). Targeting in social networks with anonymized information. *Games and Economic Behavior 141*, 380–402.

Bock, O., I. Baetge, and A. Nicklisch (2014). hroot: Hamburg Registration and Organization Online Tool. *European Economic Review 71*, 117–120.

Bramoullé, Y., A. Galeotti, and B. W. Rogers (Eds.) (2016). *The Oxford Handbook of the Economics of Networks*. Oxford University Press.

Bramoullé, Y. and R. Kranton (2007). Public goods in networks. *Journal of Economic Theory 135(1)*, 478–494.

Breza, E. (2016). Field experiments, social networks, and development. In Y. Bramoullé, A. Galeotti, and B. W. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, pp. 412–439. Oxford University Press.

Cabrales, A., D. Gale, and P. Gottardi (2016). Financial contagion in networks. In Y. Bramoullé, A. Galeotti, and B. W. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, pp. 543–568. Oxford University Press.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009). Peer effects and social networks in education. *Review of Economic Studies 76*, 1239–1267.

Chandrasekhar, A. G., H. Larreguy, and J. P. Xandri (2020). Testing models of social learning on networks: Evidence from a lab experiment in the field. *Econometrica 88(1)*, 1–32.

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

Chatterjee, K. and B. Dutta (2016). Credibility and strategic learning in networks. *International Economic Review 57(3)*, 759–786.

Chen, R. and Y. Chen (2011). The potential of social identity for equilibrium selection. *American Economic Review 101*(6), 2562–2589.

Chen, Y., S. X. Li, T. X. Liu, and M. Shih (2014). Which hat to wear? Impact of natural identities on coordination and cooperation. *Games and Economic Behavior 84*, 58–86.

Choi, S., E. Gallo, and S. Kariv (2016). Networks in the laboratory. In Y. Bramoullé, A. Galeotti, and B. W. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, pp. 440–475. Oxford University Press.

Chowdhury, S. M., D. Kovenock, and R. M. Sheremeta (2013). An experimental investigation of Colonel Blotto games. *Economic Theory 52*, 833–861.

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

Currarini, S., M. O. Jackson, and P. Pin (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica 77*(4), 1003–1045.

Currarini, S. and F. Mengel (2016). Identity, homophily and in-group bias. *European Economic Review 90*, 40–55.

de Marti, J. and Y. Zenou (2015). Networks games under incomplete information. *Journal of Mathematical Economics 61*, 221–240.

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

Demange, G. (2018). Contagion in financial networks: a threat index. *Management Science 64(2)*, 955–970.

Dixit, A. and J. Londregan (1996). The determinants of success of special interests in redistributive politics. *Journal of Politics 58(4)*, 1132–1155.

Domingos, P. and M. Richardson (2001). Mining the network value of customers. *Seventh International Conference on Knowledge Discovery and Data Mining*, 57–66.

Duarte, R., F. Finan, H. Larreguy, and L. Schechter (2025). Brokering votes with information spread via social networks. *Review of Economic Studies rdaf075, https://doi.org/10.1093/restud/rdaf075.*

Fafchamps, M. and J. Labonne (2020). Family networks and distributive politics. *Journal of the European Economic Association 18(4)*, 1697–1725.

Finan, F. and L. Schechter (2012). Vote-buying and reciprocity. *Econometrica 80(2)*, 863–881.

Friedman, D., J. Kovářík, and F. Mengel (2024). Influence in social networks. *SSRN 5059047.*

Galeotti, A. and S. Goyal (2009). Influencing the influencers: a theory of strategic diffusion. *The RAND Journal of Economics 40(3)*, 509–532.

Galeotti, A., S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv (2010). Network games. *Review of Economic Studies 77(1)*, 218–244.

Gallo, E. and C. Yan (2023). Efficiency and equilibrium in network games: An experiment. *The Review of Economics and Statistics 105(6)*, 1515–1529.

Goldenberg, J., B. Libai, and E. Muller (2001). Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters 12*, 211–223.

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

Golub, B. and E. Sadler (2016). Learning in social networks. In Y. Bramoullé, A. Galeotti, and B. W. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, Chapter 19, pp. 504–542. Oxford University Press.

Goyal, S., H. Heidari, and M. Kearns (2019). Competitive contagion in networks. *Games and Economic Behavior 113*, 58–79.

Grabisch, M., A. Mandel, A. Rusinowska, and E. Tanimura (2018). Strategic influence in social networks. *Mathematics of Operations Research 43(1)*, 29–50.

Hahn, Y., A. Islam, E. Patacchini, and Y. Zenou (2015). Teams, organization and education outcomes: Evidence from a field experiment in Bangladesh. Available at SSRN: https://ssrn.com/abstract=2613028.

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

Jackson, M. O. and L. Yariv (2011). Diffusion, strategic interaction, and social structure. In J. Benhabib, A. Bisin, and M. O. Jackson (Eds.), *Handbook of Social Economics Volume 1A*, pp. 645–678. Amsterdam: Elsevier Science.

Kempe, D., J. Kleinberg, and E. Tardos (2003). Maximizing the spread of influence through a social network. *Proceedings of 9th International Conference on Knowledge Discovery and Data Mining*, 137–146.

Kempe, D., J. Kleinberg, and E. Tardos (2005). Influential nodes in a diffusion model for social networks. *Proceedings of 32nd International Colloquium on Automata, Languages and Programming*, 1127–1138.

König, M. D., X. Liu, and Y. Zenou (2014). R&D networks: Theory, empirics and policy implications. *Review of Economics and Statistics 101(3)*, 476–491.

König, M. D., C. Tessone, and Y. Zenou (2014). Nestedness in networks: A theoretical model and some applications. *Theoretical Economics 9*, 695–752.

Kovenock, D. and B. Roberson (2021). Generalizations of the General Lotto and Colonel Blotto games. *Economic Theory 71*, 997–1032.

Lamberson, P. J. (2016). Diffusion in networks. In Y. Bramoullé, A. Galeotti, and B. W. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, pp. 479–503. Oxford University Press.

Liu, X., E. Patacchini, Y. Zenou, and L.-F. Lee (2012). Criminal networks: Who is the key player? Available at SSRN: https://ssrn.com/abstract=2089267 or http://dx.doi.org/10.2139/ssrn.2089267.

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology 27*(1), 415–444.

Nagel, R. (1995). Yunraveling in guessing games: An experimental study. *American Economic Review 85*, 1313–1326.

Nichter, S. (2008). Vote buying or turnout buying? Machine politics and the secret ballot. *American Political Science Review 102(1)*, 19–31.

Ravanilla, N., D. Haim, and A. Hicken (2022). Brokers, social networks, reciprocity, and clientelism. *American Journal of Political Science 66(4)*, 795 – 812.

Richardson, M. and P. Domingos (2002). Mining knowledge-sharing sites for viral marketing. *Eighth International Conference on Knowledge Discovery and Data Mining*, 61–70.

Roberson, B. (2006). The Colonel Blotto game. *Economic Theory 29*, 1–24.

Rosenkranz, S. and U. Weitzel (2012). Network structure and strategic investments: An experimental analysis. *Games and Economic Behavior 75*, 898–920.

Smith, R. A. and C. J. Carpenter (2018). Who persuades who? An analysis of persuasion choices related to antibiotic-free food. *Health Communication 33(4)*, 478–488.

Tsakas, N. (2017). Diffusion by imitation: The importance of targeting agents. *Journal of Economic Behavior and Organization 139*, 118–151.

Zenou, Y. (2016). Key players. In Y. Bramoullé, A. Galeotti, and B. W. Rogers (Eds.), *The Oxford Handbook of the Economics of Networks*, pp. 244–274. Oxford University Press.

A Online Appendix: Theoretical framework

A.1 Model

Our theoretical framework builds on Grabisch et al. (2018) who extended the DeGroot updating mechanism (DeGroot, 1974) to a targeting framework with two external players competing for influence in a network. We briefly recall this model.

DeGroot setting In a DeGroot setting, a set $N = \{1, \ldots, n\}$ of *non-strategic players* interact in a social network. Each player is characterized by an initial opinion on a certain issue represented by a number $x_i(0) \in [0, 1]$ and updates his opinion at discrete time instances as in DeGroot (1974), *i.e.*, by taking a convex combination of opinions of his neighbors. More precisely, there is a $n \times n$ row-stochastic *interaction matrix* of weights $W = [w_{ij}]$, where w_{ij} denotes the weight that player i places on the current opinion of player j in forming his own opinion in the next period. The evolution of opinions can be expressed as

$$\mathbf{x}_N(t+1) = W\mathbf{x}_N(t) \quad \text{for all } t \geq 0$$

where $\mathbf{x}_N(t) = [x_1(t), \ldots, x_n(t)]'$ is the (column) opinion vector[43] at time step t.

Following DeGroot (1974), Grabisch et al. (2018) associate with the matrix W a directed graph Γ on N such that there is a directed link (i, j) from i to j, meaning that i listens to the opinion of j if and only if $w_{ij} > 0$. W is assumed to be irreducible.

Next, Grabisch et al. (2018) consider two *strategic players*: a_1 with a fixed opinion 1 and a degree of influence $\mu > 0$, and a_2 with a fixed opinion 0 and a degree of influence $\lambda > 0$.[44] The authors define a non-cooperative game $\mathcal{G}_{\mu,\lambda}$ played by players a_1 and a_2. The strategy set for each of the two players is N. More precisely, each strategic player a_i chooses a strategy s_i in N that represents the non-strategic player he decides to form a link with to influence the opinion formation in the network. When a_1 targets the non-strategic player s_1, a share proportional to μ of s_1's attention is redirected to a_1. Similarly, when a_2 targets s_2, a share proportional to λ of s_2's attention is redirected to a_2. Each strategic player aims to bring the asymptotic average opinion in the network as close as possible to his own opinion (1 for a_1 and 0 for a_2).

[43]We denote by $'$ the transposition of vectors.

[44]In our experimental design, we fix $\mu = 1$.

With the ordering of players $a_1, a_2, 1, \ldots, n$ and the choice of targets $\mathbf{s} = (s_1, s_2)$, the $n \times n$ matrix W is extended to a $(n+2) \times (n+2)$ matrix $P_{\mu,\lambda}(\mathbf{s})$ given by:

$$P_{\mu,\lambda}(\mathbf{s}) = \left[\begin{array}{cc|c} 1 & 0 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \\ \hline \multicolumn{2}{c|}{R_{\mu,\lambda}(\mathbf{s})} & Q_{\mu,\lambda}(\mathbf{s}) \end{array}\right] \tag{3}$$

where

$$R_{\mu,\lambda}(\mathbf{s}) = \Delta_{\mu,\lambda}(\mathbf{s}) E_{\mu,\lambda}(\mathbf{s}), \quad Q_{\mu,\lambda}(\mathbf{s}) = \Delta_{\mu,\lambda}(\mathbf{s}) W \tag{4}$$

The weight renormalization matrix $\Delta_{\mu,\lambda}(\mathbf{s})$ is the diagonal matrix with diagonal elements

$$\frac{d_1}{d_1 + \mu\delta_{1,s_1} + \lambda\delta_{1,s_2}}, \ldots, \frac{d_n}{d_n + \mu\delta_{n,s_1} + \lambda\delta_{n,s_2}}$$

where d_i is the number of outgoing links of $i \in N$ and δ is the Kronecker symbol, i.e., $\delta_{i,s_j} = 1$ if $i = s_j$ and 0 otherwise, and the matrix $E_{\mu,\lambda}(\mathbf{s})$ is equal to

$$E_{\mu,\lambda}(\mathbf{s}) = \begin{bmatrix} \dfrac{\mu}{d_{s_1}} e_{s_1} & \dfrac{\lambda}{d_{s_2}} e_{s_2} \end{bmatrix}$$

where e_i denotes the unit vector with coordinate 1 at i.

In other words, the matrix $P_{\mu,\lambda}(\mathbf{s})$ is the $(n+2) \times (n+2)$ extended interaction matrix of weights, where each strategic player is stubborn and puts the weight 1 to himself, while every non-strategic player recalculates and assigns the weights to all his neighbors, including the strategic player(s) if he has been targeted.

The opinion vector in time t is extended to a $(n+2)$-vector $\mathbf{x}(t)$ and the opinion updating is given by:

$$\mathbf{x}(t+1) = P_{\mu,\lambda}(\mathbf{s})\mathbf{x}(t) \tag{5}$$

i.e., the opinions of the strategic players are constant and equal to 1 and 0, respectively, and the opinions of the non-strategic players are updated as follows:

$$\mathbf{x}_N(t+1) = \Delta_{\mu,\lambda}(\mathbf{s}) E_{\mu,\lambda}(\mathbf{s}) \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \Delta_{\mu,\lambda}(\mathbf{s}) W \mathbf{x}_N(t) \tag{6}$$

Note that the evolution law for the opinions of the non-strategic players determined in (6) is simply obtained by inserting the matrices $P_{\mu,\lambda}(\mathbf{s})$, $R_{\mu,\lambda}(\mathbf{s})$ and $Q_{\mu,\lambda}(\mathbf{s})$ defined in (3) and (4) into the law of motion given in (5).

Intermediacy and influenceability Finally, Grabisch et al. (2018) introduce two measures useful in characterizing the equilibrium of the game: intermediacy and influenceability. They are defined by how one node can be reached from another, *i.e.*, how influence spreads in a network. Moreover, assuming that the influence spreads according to the probabilities specified by W, these measures have a natural probabilistic interpretation.

Intermediacy b^i_j of player i relatively to player j is equal to the sum of weights of walks[45] to j that pass through i. It measures the extent to which the influence of i reaches the network before that of j. Equivalently, b^i_j can be interpreted as the sum of the probabilities for all players distinct from j to be reached by the influence of i before that of j. The *intermediacy centrality* B_i of player i is equal to the minimal sum of weights of the walks to a given node passing through i, *i.e.*,

$$B_i = \min_{j \neq i} b^i_j \tag{7}$$

In other words, a player with a high intermediacy centrality must maximize the minimal influence with respect to any other player in the network.[46]

Influenceability of player i, given that player j is targeted by the other strategic player, is measured by $d_i c^j_i$, where d_i is the number of the outgoing links of i and c^j_i is the sum of weights of cycles around i that pass through j.[47] Also, c^j_i can be interpreted as the probability for i to be reached by the influence of j before he gets the self-feedback of his own opinion. This is a decreasing measure, *i.e.*, the larger $d_i c^j_i$, the less influenceable i is. Indeed, the larger d_i, the more opinions i considers, and the slower i is influenced by an additional opinion. Also, the larger c^j_i, the lesser the influence that can be exerted on i by a strategic player.

[45] A walk in Γ between nodes i and j is a sequence of directed links $(i_1, i_2), \ldots, (i_{K-1}, i_K)$ such that every (i_k, i_{k+1}) belongs to Γ for $k \in \{1, \ldots, K-1\}$, $i_1 = i$ and $i_K = j$. The weight of such a walk measured according to W is equal to the multiplication of the weights of all directed links in the walk, *i.e.*, $\prod_{k=1}^{K-1} w_{i_k, i_{k+1}}$.

[46] Note that due to its strategic nature, the intermediacy centrality differs from other established centrality measures. See Appendix C in Grabisch et al. (2018) for examples.

[47] A walk from i to i which does not pass through i between the starting and the ending points is called a cycle around i.

A.2 Convergence and equilibria

We now summarize some theoretical results of Grabisch et al. (2018) and add some remarks that interest our experimental study. In a non-competitive framework, every non-strategic player is an optimal target and the targeting choice only affects the speed of convergence. It is different when there are multiple strategic players.

Convergence The vector of the asymptotic opinions $\overline{\mathbf{x}}_N \in [0,1]^n$ of the non-strategic players is given by:

$$\overline{\mathbf{x}}_N = (I - \Delta_{\mu,\lambda}(\mathbf{s})W)^{-1}\frac{\mu}{d_{s_1}}\Delta_{\mu,\lambda}(\mathbf{s})e_{s_1}$$

In contrast to the DeGroot model, the asymptotic opinions are independent of the initial opinion vector. This results from the introduction of two strategic players who influence the opinions of non-strategic players, but remain stubborn and never alter their own opinion. The asymptotic opinions instead are determined by the targeting choices of the strategic agents and their degrees of influence and the network structure, as the steady-state vector $\overline{\mathbf{x}}_N$ depends on the strategy vector $\mathbf{s}$, the interaction matrix W, μ and λ. These are the dimensions that we have to manipulate in our experiment to study targeting decisions.

Nash equilibria of the game The game $\mathcal{G}_{\mu,\lambda}$ is a constant-sum game played by the strategic players a_1 and a_2, with payoffs equal to:

$$\pi_{\mu,\lambda}(\mathbf{s}) = \mathbf{1}' \cdot \overline{\mathbf{x}}_N = \mathbf{1}' \cdot (I - \Delta_{\mu,\lambda}(\mathbf{s})W)^{-1}\frac{\mu}{d_{s_1}}\Delta_{\mu,\lambda}(\mathbf{s})e_{s_1}$$

and $n - \pi_{\mu,\lambda}(\mathbf{s})$ for a_1 and a_2, respectively. Equilibria differ depending on whether the two strategic players have the same level of influence or not.

Grabisch et al. (2018) characterize equilibria in pure strategies with equal levels of influence ($\mu = \lambda$). A pair of strategies (i,i) is an equilibrium of the game $\mathcal{G}_\mu$ if and only if for all $j \in N \setminus \{i\}$:

$$\mu\left[b_j^i - b_i^j\right] \geq n\left[d_i c_i^j - d_j c_j^i\right] \tag{8}$$

In other words, (i,i) is an equilibrium if for every $j \neq i$, the excess intermediacy of i over j is not smaller than the excess influenceability of j over i, scaled by the factor $\frac{n}{\mu}$.

The existence and uniqueness of a pure strategy equilibrium depend on the structure of the network. There are networks for which there exist equilibria in pure strategies and networks for which there exists no equilibrium in pure strategies.

The relative importance of intermediacy versus influenceability increases with the level of influence μ. The strategy profile (i, i) is a Nash equilibrium of $\mathcal{G}_\mu$ for any arbitrarily high degree of influence μ only if the relative intermediacy of player i exceeds that of any other player j. On the other hand, the strategy profile (i, i) is a Nash equilibrium of $\mathcal{G}_\mu$ for a vanishingly low degree of influence μ only if i is more influenceable than any other player j.

Concerning non-symmetric equilibria, if a pair of strategies (i, j) is an equilibrium of the game $\mathcal{G}_\mu$, then also (j, i), (i, i) and (j, j) are equilibria.

For the case of unequal degrees of influence ($\mu \neq \lambda$), if the difference between the players' influence degrees is sufficiently large, the game has only equilibria in mixed strategies. Moreover, with different degrees of influence (without loss of generality, let $\mu < \lambda$), the relative importance of intermediacy *vs.* influenceability increases with the magnitude of the players' influence. Consider $\mu > 0$ as fixed (as in the experiment). When λ increases, player a_1 should respond asymmetrically, as he or she is better off when being the sole player targeting a given node. As player a_2 focuses on intermediacy, a_1 should target more influenceable players. This is a key prediction of the theoretical setting that we test in our experiment.

B Online Appendix: Instructions and questionnaires [*Translated from French*]

B.1 Instructions [*Common to all treatments*]

Welcome!

Thank you for participating in this experimental session on decision-making. Please turn off your phone and put it away. You are not allowed to communicate with other participants, or you will be disqualified from the session and earnings.

All the decisions you make are anonymous.

The experiment consists of three independent parts. You will receive instructions for the next part once the previous part is finished.

You will earn 5 Euros for showing up on time. In addition, you can accumulate earnings in each part based on your decisions. Transactions are expressed in Experimental Currency Units (ECU), convertible into Euros at the rate of: 110 ECU = 1 Euro.

At the end of the session, your earnings in Euros will be paid in cash in a separate room and in private. Your earnings will remain confidential.

Part 1

In this part, you interact with all other participants in this session.

Everyone must choose a number between 0 and 100. The winner will be the one of you whose chosen number is closest to two-thirds (2/3) of the average number chosen by all participants.

The winner earns **1000 ECU (€10)**. The other participants do not earn anything. In the event of a tie, earnings will be shared equally between the winners. You will be informed whether you are a winner or not at the end of the session.

If you have any questions regarding the instructions at any time during the session, please raise your hand or press the red button on the side of your desk. We will come and answer your questions in private.

—

Part 2 [*Distributed after completion of part 1*]

This part consists of **8 subparts of 5 periods each**.

The roles

In this part, there are two roles: A and B. There are 2 participants A and 5 participants B. **You have the role of participant A**. You will always keep the same role.

The participants B are linked together in a network whose shape can vary during the experiment.

The other participant A and the 5 participants B with whom you are matched are not real people: they are **virtual participants** represented by the computer program.

You and the other participant A have opposing views on a topic (you can think of a political, economic, or social topic). The other participant A is referred to as **your Opponent**.

Throughout the part your opinion is always 1 and your Opponent's opinion is always 0. You and your Opponent never change your opinion during the whole part.

In some cases, **all participants B have the same opinion** at the beginning of the period: **0.5** (they are indifferent). In some other cases, **participants B have different opinions** at the beginning of the period (**0.25, 0.5 or 0.75**): for each participant B, each value ex-ante has a probability of being chosen of 33.3% but we impose that each opinion is represented at least once in each network.

In each case, participants B's opinions change throughout the period based on their links to you, your Opponent, and the other B participants, as explained below.

Your task

Your task in each period is **to choose a single participant B** so that the average opinion of participants B is as close as possible to your opinion.

Your Opponent also chooses a participant B intending to get the average opinion of participants B as close as possible to his or her opinion.

You are informed of your Opponent's choice before you make your own choice. You may or may not choose the same B participant as your Opponent and you may or may not change it from one period to the next.

By making these choices, you and your Opponent change **the weights of the opinions** of each participant to whom a participant B is linked, according to your degree of influence.

The weight of the opinions among participants

Each participant B is linked to one or more other Bs. These links determine the weight of the opinion of each of the others linked to him or her on his or her own opinion.

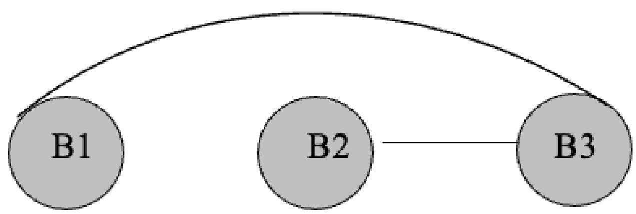
The weight of an opinion indicates the relative importance given by a participant B to the opinion of the participant with whom he or she has **a direct link**.

For each B, **the sum of the weights of the opinions** of those to which he or she is **directly** linked is **always equal to 1**. Thus, the higher the number of links, the lower the weight of each.

Before participants A make their decision, if a participant B is not linked to another B, the weight of the opinion of that other B is 0.

- If he or she has a link with only 1 other B, the weight of the opinion of this other B is 1.
- If he or she has a link with 2 other Bs, the weight of the opinion of each of the 2 other Bs is 0.5.
- If he or she has a link with 3 other Bs, the weight of the opinion of each of the 3 other Bs is 0.33.
- If he or she has a link with 4 other Bs, the weight of the opinion of each of the 4 other Bs is 0.25.

Please consider this simplified example with only three Bs:



In this example, for B1 the weight of B2's opinion is 0 (since he or she is not linked to him or her) and that of B3's opinion is 1.

For B2, the weight of B3's opinion is 1 and the weight of B1's opinion is 0.

For B3, the weight of B1's option is 0.5 and that of B2's opinion is 0.5 (since he or she is related to the other two Bs).

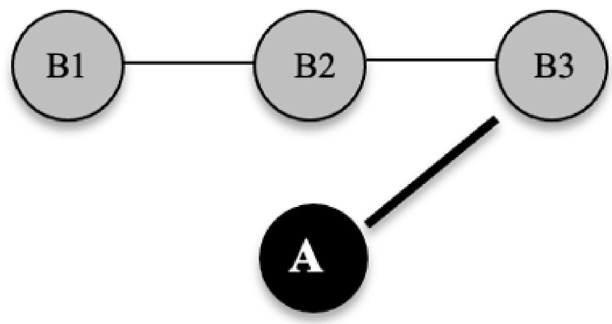
Degree of influence of participants A

Your degree of influence is always equal to 1. It never changes during the part. This means that your opinion counts as much for the chosen participant B as the opinion of every other B he or she is linked to.

The degree of influence of your Opponent is equal to X, with X varying from 1 to 5 depending on the period. This means that his or her opinion counts X times more for your Opponent's chosen participant B than the opinion of each of the other Bs he or she is linked to. For example, if X=5, your Opponent's opinion counts 5 times more for the participant he or she chose than each of the other links for that participant B.

Suppose that your Opponent establishes a link with a given participant B. Then the weights of the opinions of each of the links of this participant B change, as explained before, but with a greater weight for the opinion of your Opponent if his or her degree of influence is greater than 1.

Please consider this simplified example with only three Bs and one A:



In this example, before A created a link, the weight of B2's opinion on B3 was 1.

After the creation of the link by A and if the degree of influence of A is 1, the weight of the opinion of B2 on B3 becomes 0.5 and that of the opinion of A on B3 is 0.5.

If the degree of influence of A is 3, the weight of the opinion of B2 on B3 becomes 0.25 and that of the opinion of A on B3 is 0.75.

Outline of the periods

A black avatar represents you. Your opinion is always equal to 1. A white avatar represents Your Opponent. His or her opinion is always equal to 0.

A graph represents the links between the 5 participants B. Participants B are represented by a grey circle and a number. Their opinion (initially 0.25, 0.5, or 0.75) is indicated next to the circle. The color of the circle is more or less dark according to their opinion.

At the beginning of each of the 8 subparts (counting 5 periods each), you are informed of:

- the links between participants B
- the initial opinions of participants B
- the participant B chosen by your Opponent.

These links and your Opponent's choice remain the same within a subpart but may change between subparts.

At the beginning of each of the 5 periods, before making your choice, you are informed of:

- the degree of influence of your Opponent, which may change from one period to another.

Outline of each period

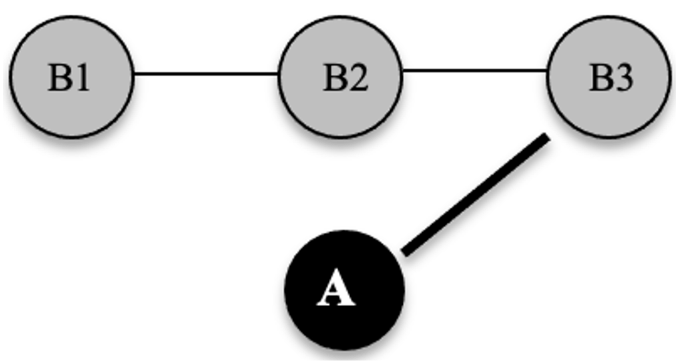
After having consulted the degree of influence of your Opponent, **you must choose one of the five participants B** by clicking on the grey circle representing him or her. You can choose the same participant as your Opponent.

The screenshot below represents a situation in which all B participants have an initial opinion of 0.5, your Opponent chose B3 and you choose B4:

Once you and your Opponent made your choices, the weights of opinions stay fixed for the period. The program calculates the final opinions of B participants for the period based on an iterative adjustment process explained as below.

The program calculates for each B **the sum of the opinions of the participants to which he or she is linked, weighted by their respective weights.**

Let's take the previous simplified example with only one player A and 3 players B:



Suppose that A's degree of influence is 1 and that A chooses B3 which is also linked to B2. The initial opinion of B is 0.5 and that of A is 1. B3 now gives a weight of 0.5 to A's opinion and 0.5 to B2's opinion. The opinion of B3 becomes:

*(1 (i.e., the opinion of A) * 0.5 (i.e., the weight of A))*

*+(0.5(i.e., the opinion of B2) * 0.5 (i.e., the weight of B2)) = 0.75.*

This triggers **an iterative process of revision of opinions** as each person's new opinion influences the opinions of those to whom he or she is linked.

The opinion of each B is influenced by the A and B participants to whom he or she is directly linked, as well as by the participants to whom he or she is indirectly connected through the direct links.

In the previous simplified example, B2's opinion is also revised by taking into account B3's new opinion, which at the same time influences the opinions of the other B participants linked to B2. This dynamic process can be described with the following iterations:

*Specifically, B2's opinion was 0.5 (= opinion of B1) * 0.5 (= weight of B1) + 0.5 (= opinion of B3) * 0.5 (= weight B3) = 0.5. It becomes: 0.5 (= opinion of B1) * 0.5 (=weight of B1) + 0.75 (= new opinion of B3) * 0.5 (= weight of B3) = 0.625.*

*B1's opinion is also affected. It was 0.5 (= opinion of B2) * 1 (= weight of B2) = 0.5. It becomes: 0.625 (=new opinion of B2) *1 (= weight of B2) = 0.625.*

*Simultaneously, B3 revises his opinion in reaction to B2's change. It becomes: 0.625 (= new opinion of B2) *0.5 + 1 (= opinion of A) * 0.5 = 0.812.*

*B2 also revises his opinion. It becomes: 0.625 (= new opinion of B1) * 0.5 + 0.812 (= new opinion of B3) * 0.5 = 0.719.*

All B participants revise their opinions simultaneously in each iteration. The opinion revision process continues **until all B opinions stabilize.**

On your screen, you can observe the continuous updating of the Bs' opinions until they stabilize. At the same time, **the color of the Bs changes**: it turns white if the opinion is closer to your Opponent's; it turns black if it is closer to your opinion.

Once opinions are stabilized, **you observe how close the average opinion of the Bs has come to your opinion or to that of your Opponent.**

End of each period and sequence of periods

At the end of the opinion revision process, you observe:

- the final opinion of each participant B
- the average final opinion
- your potential earnings in case this period would be selected randomly for the final payment

Then a new period or a new subpart will automatically follow.

As a reminder:

- **At the beginning of each new subpart**, the links between the B participants and the B participant chosen by your Opponent may change. These links and your Opponent's choice remain the same within the same subpart.
- **At the beginning of each new period**, the degree of influence of your Opponent changes.

Calculation of earnings for the period and the part

Your gain for the period is **higher the closer the average final opinion of B participants was to yours.**

You earn **600 ECU** if you achieved the maximum possible alignment of the average opinion of the B participants with your own opinion, considering the initial links and the degree of influence of your Opponent.

Your earning is reduced by one percentage point (thus, 6 ECU) for every one percentage point deviation from this maximum. For example, if you have obtained 95% of the maximum, you earn 95% of 600 ECU (570 ECU); with 62% of the maximum, you earn 62% of 600 ECU (372 ECU), etc.

At the end of the session, the program will draw 3 sub-parts and, in each sub-part, one of the 5 periods. Your earnings in these 3 periods will be added together and will constitute your earnings for the part.

—

Please read again these instructions. If you have any questions, please raise your hand or press the red button and we will answer your questions in private.

Once this is complete, you can answer questions on your computer to ensure that you understand the instructions.

Next, you can practice manipulating the interface for 6 minutes. You can change the degree of influence of your Opponent, the participant B chosen by your Opponent, your choice of participant B. For 3 minutes, all the B participants have the same initial opinion, and for 3 other minutes, the B participants do not necessarily have the same initial opinion.

—

Vocabulary – Reminder

The weight of an opinion indicates the relative importance given by a participant B to the opinion of the participant with whom he or she has a direct link. The sum of the weights of opinions of whom he or she is directly linked is always equal to 1.

The degree of influence of your Opponent is equal to X (X varying **from 1 to 5**): his or her opinion counts X times for the participant B chosen by your Opponent than the opinion of each of the other B participants to whom he or she is linked.

—

Part 3 [***On screen, after completion of part 2***]

In this part, you have to answer 6 questions.

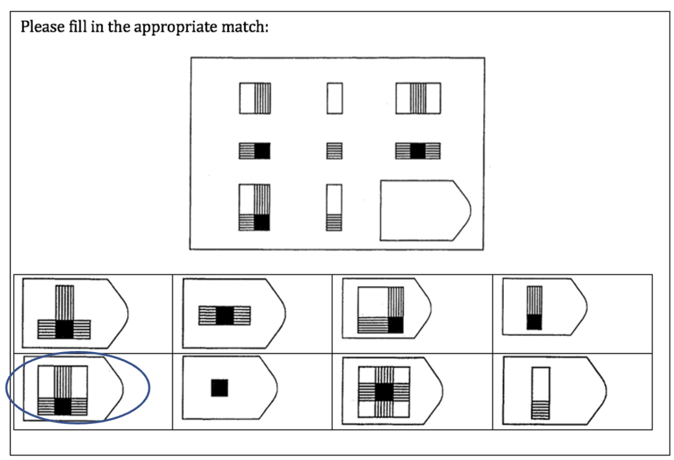
For each question, a series of figures will appear. Your task is to identify which figure logically follows the previous ones.

For each question, 8 possible figures are suggested. You have to tick the right answer among these 8 figures then validate your choice.

You have 6 minutes to answer the 6 questions.

You earn 50 ECU for each correct answer. You lose nothing for an incorrect answer.

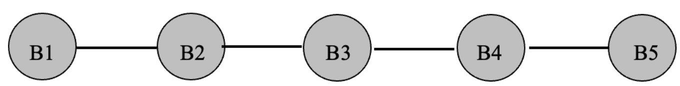
In the example below, the correct answer is circled.



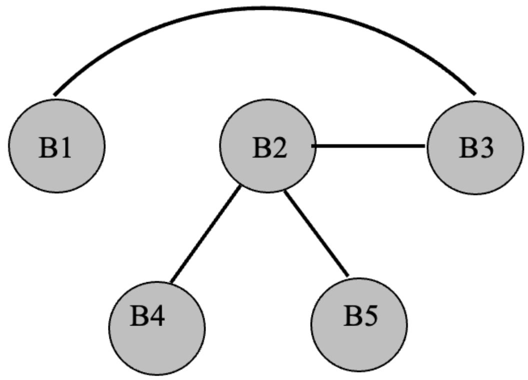
B.2 Comprehension questionnaire [*Displayed on screen at the beginning of part 2*]

Please answer the following questions.

- Your Opponent always chooses the same participant B between periods of a subpart: T/F *(T: Your Opponent always chooses the same participant B in a block)*
- For a given participant B, the sum of the weights of the other participants is always equal to 1: T/F *(T: The sum of the weights of the other participants is always equal to 1)*
- For a given participant B, the weight of the opinions of each other participant cannot change during a period: T/F *(F: The relative weight of the opinions of participants B changes according to your choice of link and that of your Opponent)*
- Between periods, your degree of influence changes: T/F *(F: Your degree of influence is always 1; it is your Opponent's degree of influence that changes according to the periods)*
- Your opinion is always equal to 1 and your Opponent's to 0: T/F *(T: Your opinion and that of your Opponent never changes)*
- In the following initial configuration:

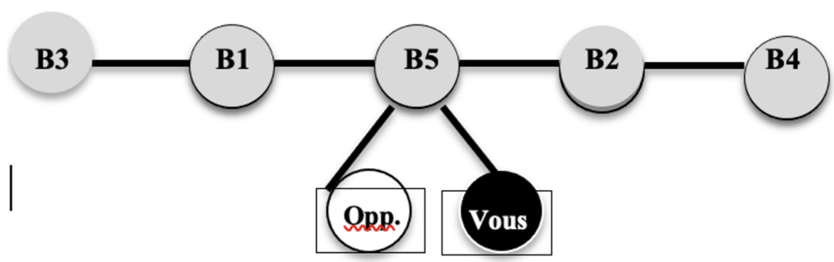


  - What is the weight of B2's opinion on B1? *(1)*
  - What is the weight of B2's opinion on B3?*(0.5)*
  - What is the weight of B4's opinion on B3? *(0.5)*
  - What is the weight of B4's opinion on B5? *(1)*
- In the following initial configuration:



- What is the weight of B2's opinion on B1? *(0)*
- What is the weight of B3's opinion on B1? *(1)*
- What is the weight of B5's opinion on B2? *(0.33)*
- What is the weight of B1's opinion on B3? *(0.5)*

• Suppose that your Opponent's degree of influence is 2. Your Opponent chose B5 which is linked to B1 and B2. You have also chosen B5.



- What is the weight of your Opponent's opinion on B5? *(0.4)*
- What is the weight of your opinion on B5? *(0.2)*
- What is the weight of B1's opinion on B5? *(0.2)*
- What is the weight of B2's opinion on B5? *(0.2)*

(If your Opponent has a degree of influence of 2, his or her opinion carries twice the weight of any of the other links in B5. Thus, B5 gives a weight of 0.4 to your Opponent's opinion, a weight of 0.2 each to your opinion, B1's and B2's.)

B.3 Final questionnaire [*Displayed on screen after the completion of part 3*]

Please, answer the following questions.

- How old are you?
- What is your gender?
  - Female
  - Male
- What is your level of education?
  - Without diploma
  - Certificat des écoles
  - Brevet
  - BEP - CAP
  - Baccalaureat
  - Bac +2 - DEUG - IUT - DUT - BTS
  - Bac +3 - Licence
  - Bac +4 - Maitrise
  - Bac +5 - Master - DESS - DEA
  - Higher than Bac +5 - Doctorat - Thèse
- To which socio-professional category do you belong?
  - Student
  - Employee
  - Unemployed
  - Inactive (homemaker, pensioner, disabled
- What is your school if you are a student? ("Other" if you are not a student)
  - ECLyon
  - EMLyon
  - University

  - ITECH
  - ISOSTEO
  - Other

- What was your average Baccalaureat grade? (*between 10 and 20*)
- How much do you spend in a typical week (for your ordinary expenses, such as food, leisure, travel, excluding rent, loans and charges)
  - Less than €100
  - €101 - €200
  - More than €200
- In general, please indicate to what extent are you willing to take risks on a scale going from 0 to 10, where 0 means that you are "not at all willing to take risks" and 10 means that "you are very willing to take risks". You can also use any digits between 0 and 10 to indicate where you are on the scale.
- Imagine the following situation: You have unexpectedly received 1000 euros today. How much of this amount you would give to a charity? (Values between 0 and 1000 are authorized.): Euro.
- To what extent are you willing to make donations to charity without expecting anything in return? Please indicate your response on a scale going from 0 to 10. 0 means "not at all willing' and 10 means "very willing". You can also use any digits between 0 and 10 to indicate where you are on the scale.
- How much time do you spend on average on social media? (Facebook, Tiktok, Twitter, etc.)?
  - Never
  - Every week but not every day
  - Less than 30 minutes per day
  - Between 30 and 59 minutes per day
  - Between 60 and 119 minutes per day
  - Between 120 and 179 minutes per day
  - More than 180 minutes per day
- Do you use Twitter?

  - Yes
  - No

- If you have responded Yes, how many messages do you tweet or retweet on average in a typical day?
  - Less than 1
  - Between 1 and 4
  - Between 5 and 9
  - 10 and more

- Do you play strategy games (chess, go...)?
  - Yes
  - No

- Do you play team sports?
  - Yes
  - No

- What is the perceived degree of difficulty of the decisions in Part 2 on a scale of 0 (very simple) to 10 (very complex)?

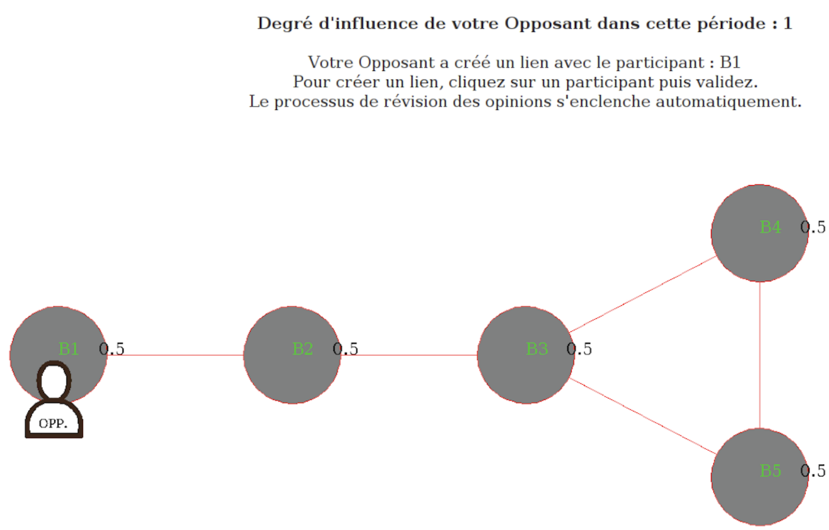
- Of your 40 decisions in Part 2, how many did you make at random?

- Please indicate to what extent you have a positive or negative feeling for each question asked. A 0 score indicates the greatest negativity and a 100 score indicates the greatest positivity. A 50 score indicates that you are feeling neutral concerning the question. It is possible to not respond to a question if you want but note that we do not keep a record of the answers but only of the aggregate indices.
  - Abortion
  - Government restrained to core functions
  - Military and national security
  - Religion
  - Social benefits
  - Traditional marriage

- Traditional values
- Fiscal responsibility
- Business world
- Family unit
- Patriotism

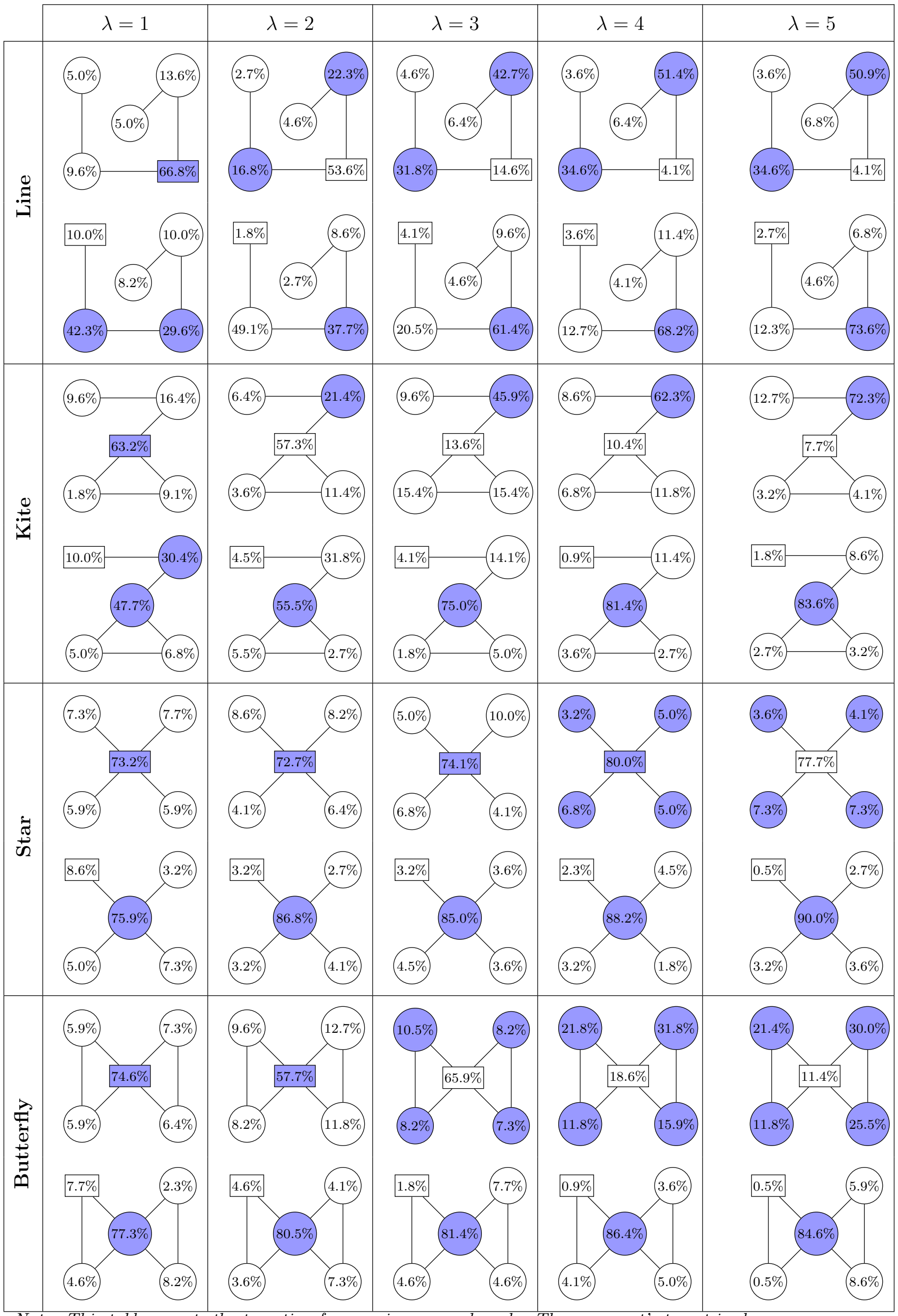
C Online Appendix: Figures

Figure C1: Example of a participant's screen



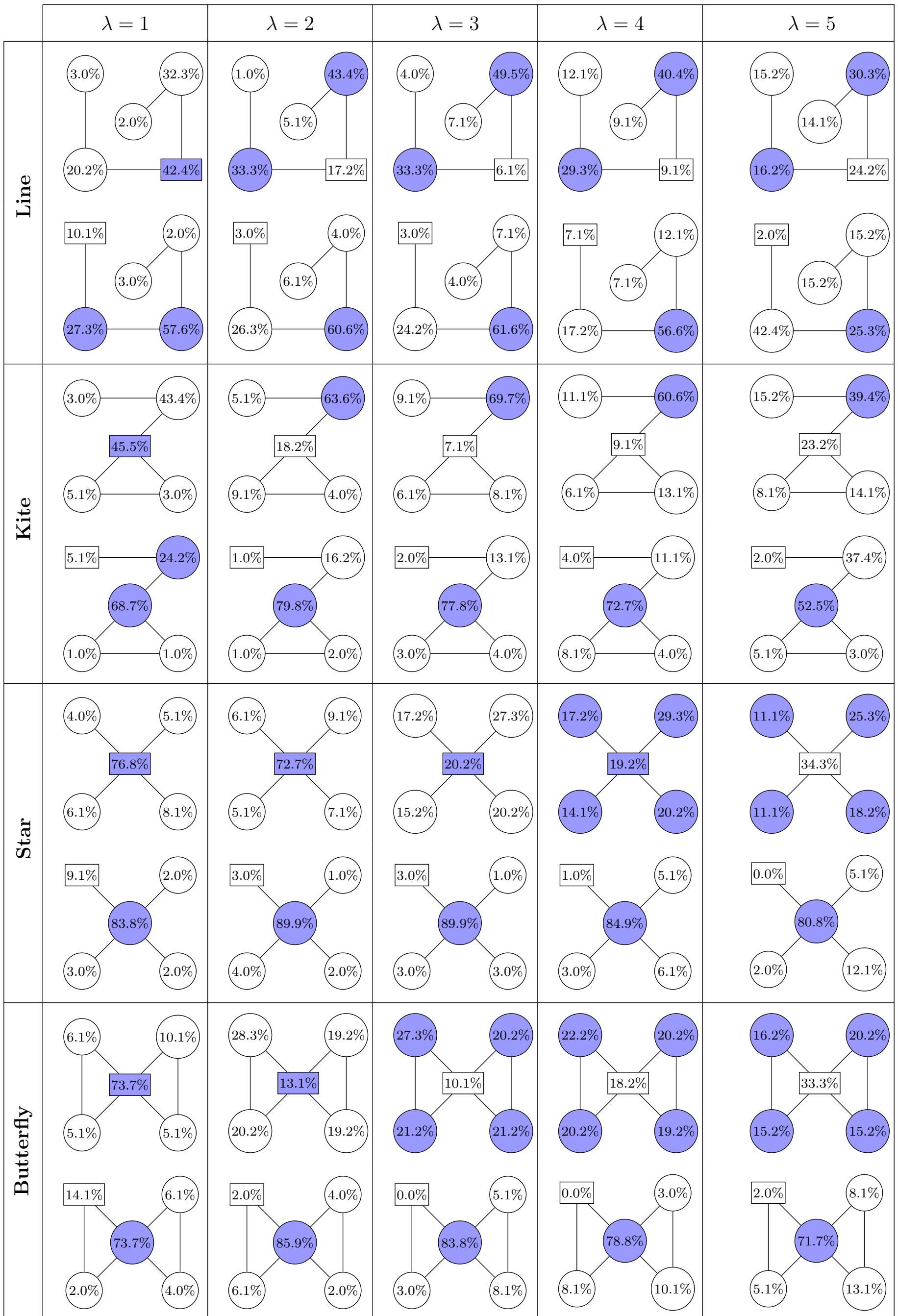
Note: The figure illustrates a participant's screen with a kite network and nodes with homogeneous initial opinions. The white avatar represents the opponent that, in this example, selected a peripheral node.

Figure C2: Relative frequencies of targeting choices, by condition - Increasing Influence treatment



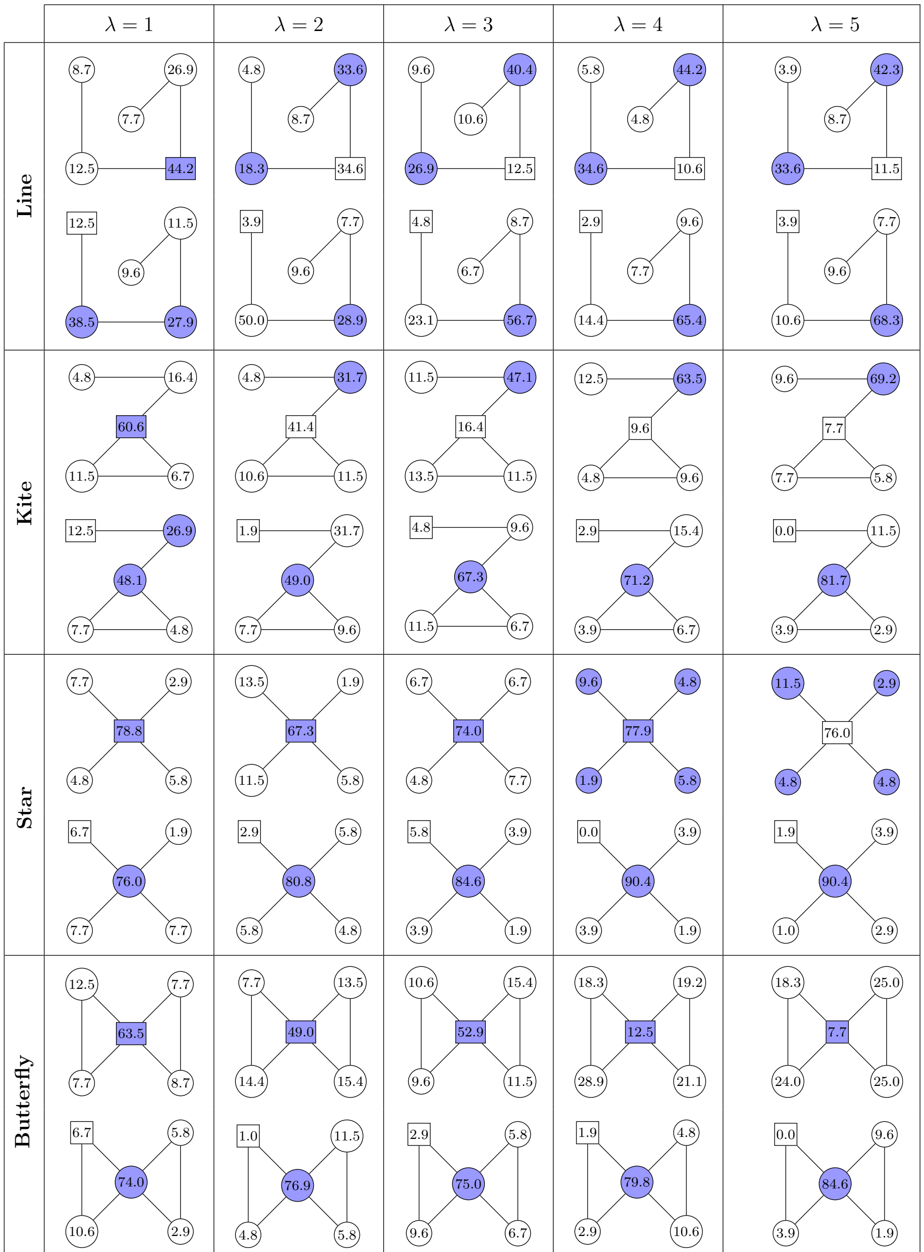
Note: This table reports the targeting frequencies per each node. The opponent's target is shown as a square, and the player's best response is indicated in blue.

Figure C3: Relative frequencies of targeting choices, by condition - Decreasing Influence treatment



Note: This figure reports the targeting frequencies per each node. The opponent's target is shown as a square, and the player's best response(s) is indicated in blue.

Figure C4: Relative frequencies of targeting choices, by condition - Layout treatment



Note: This table reports the targeting frequencies for each node. The opponent's target is shown as a square, and the player's best response(s) is indicated in blue.

D Online Appendix: Tables

Table D1: Summary statistics of participants

Treatment	(1) Increasing Influence		(2) Decreasing Influence		(3) Layout		(1 − 2) *p-value*	(1 − 3) *p-value*
	Mean	s.d.	Mean	s.d.	Mean	s.d.		
Age (Years)	21.06	1.52	22.24	3.43	21.56	1.93	0.00***	0.01**
Male (%)	0.51	0.50	0.57	0.50	0.47	0.50	0.35	0.52
Central engineering School (%)	0.41	0.49	0.51	0.50	0.35	0.48	0.11	0.28
Social media < 1H/day (%)	0.26	0.44	0.28	0.45	0.25	0.44	0.66	0.86
Risk preferences	6.44	1.88	6.41	1.76	6.58	1.68	0.71	0.68
Pro-sociality	115.75	174.91	111.07	194.95	120.30	177.25	0.13	0.57
Number of observations	220		99		104			

Notes: This table summarizes the socio-demographic characteristics of participants for each treatment. The social media variable reports the percentage of participants using social media less than 1 hour per day. The risk variable is the willingness to take risks on a scale from 0 to 10. The pro-sociality variable is the willingness to donate money on a scale from 0 to 1000. The p-values are computed from Mann-Whitney rank-sum tests. *** p<0.01, ** p<0.05, * p<0.1.

Table D2: Relative frequency of best responses, by treatment

Treatment	(1) Increasing Influence	(2) Decreasing Influence	(3) Layout	(1-2) *p-value*	(1-3) *p-value*
All data	0.70	0.69	0.68	0.204	0.279
Opponent's target					
Center	0.65	0.63	0.64	0.223	0.409
Periphery	0.76	0.74	0.72	0.294	0.226
Best response					
Center	0.75	0.69	0.71	0.001***	0.101
Periphery	0.59	0.68	0.61	0.001***	0.663
Opponent's degree of influence					
$\lambda = 1$	0.73	0.72	0.67	0.256	0.118
$\lambda = 2$	0.56	0.68	0.54	0.000***	0.634
$\lambda = 3$	0.66	0.72	0.65	0.006***	0.687
$\lambda = 4$	0.82	0.76	0.80	0.004***	0.365
$\lambda = 5$	0.75	0.56	0.73	0.000***	0.468
Network structure					
Line	0.66	0.61	0.60	0.016**	0.022**
Kite	0.64	0.65	0.62	0.632	0.469
Star	0.77	0.76	0.77	0.930	0.601
Butterfly	0.75	0.72	0.73	0.088*	0.545
Homogeneous	0.72	0.69	0.69	0.202	0.318
Heterogeneous	0.69	0.68	0.67	0.523	0.286
Unique BR	0.71	0.65	0.67	0.003	0.171
Multiple BR	0.69	0.78	0.70	0.000***	0.991
N players	220	99	104	-	-

Notes: This table summarizes the percentage of best responses played by participants, by treatment and condition. *P*-values are from Mann-Whitney rank-sum tests. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table D3: Relative frequency of best responses in the Increasing Influence treatment, with p-values

All networks	0.7					
Opponent's target	(1)	(2)	p-value			
	Center	Periphery				
	0.65	0.76	< 0.001			
Best response	(1)	(2)	(1)-(2)			
in the network	Center	Periphery	p-value			
	0.75	0.59	< 0.001			
Opponent's degree	(1)	(2)	(3)	(4)	(5)	
of influence	$\lambda = 1$	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$	$\lambda = 5$	
	0.73	0.56	0.66	0.82	0.75	
	(1)-(2)	(1)-(3)	(1)-(4)	(1)-(5)	(2)-(3)	
p-value	< 0.001	< 0.001	< 0.001	0.5278	< 0.001	
	(2)-(4)	(2)-(5)	(3)-(4)	(3)-(5)	(4)-(5)	
	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	
Network structure	(1)	(2)	(3)	(4)		
	Line	Kite	Star	Butterfly		
	0.67	0.64	0.77	0.75		
	(1)-(2)	(1)-(3)	(1)-(4)	(2)-(3)	(2)-(4)	(3)-(4)
p-value	0.088	< 0.001	< 0.001	< 0.001	< 0.001	0.057
# Best responses	(1)	(2)	(1)-(2)			
in the network	Unique	Multiple	p-value			
	0.71	0.69	0.066			
Nodes' initial	(1)	(2)	(1)-(2)			
opinion	Homog.	Heterog.	p-value			
	0.72	0.69	0.010			

Notes: P-values are from Wilcoxon signed-rank tests.

Table D4: Average marginal effects from conditional logit on the probability of targeting a node

Dependent variable: Choice of a node	(1) All blocks	(2) All blocks	(3) Heterogeneous blocks	(4) Best response is not center
Best response	0.246***	0.244***	0.242***	0.185***
	(0.007)	(0.007)	(0.008)	(0.011)
Node targeted by Opponent	-0.061***	-0.063***	-0.096***	
	(0.008)	(0.008)	(0.011)	
Node in the center	0.254***	0.256***	0.239***	
	(0.008)	(0.008)	(0.010)	
Node's initial opinion: 0.25		-0.052***	-0.049***	-0.052***
		(0.012)	(0.011)	(0.018)
Node's initial opinion: 0.75		0.102***	0.099***	0.106***
		(0.011)	(0.010)	(0.015)
Observations	84,600	84,600	42,300	25,380
Pseudo-R-squared	0.276	0.286	0.269	0.050
Log-pseudo likelihood	-19705.9	-19440.8	-9955.8	-7764.9

Notes: This table reports the average marginal effects from a conditional Logit model by decision set (participant X block X period). The effects represent the change in predicted probability when the variable switches from 0 to 1. Robust standard errors, clustered at the individual level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table D5: Conditional Logit estimates of targeting a given node - Heterogeneity analysis, part A

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Opinion Homog.	Opinion Heterog.	Treat. Increas.	Treat. Decreas.	Treat. Layout	Blocks: 1 to 4	Blocks: 5 to 8
Best Response	1.198***	1.166***	1.217***	1.181***	1.128***	1.034***	1.312***
	(0.049)	(0.044)	(0.055)	(0.082)	(0.076)	(0.050)	(0.046)
Opponent's	-0.141***	-0.463***	-0.135**	-0.749***	-0.256***	-0.286***	-0.340***
target	(0.049)	(0.051)	(0.055)	(0.079)	(0.081)	(0.049)	(0.061)
Center	1.288***	1.153***	1.291***	1.131***	1.166***	1.201***	1.279***
	(0.047)	(0.050)	(0.056)	(0.097)	(0.070)	(0.056)	(0.057)
Opinion: 0.25		-0.238***	-0.178**	-0.343***	-0.328***	-0.249***	-0.264***
		(0.054)	(0.079)	(0.115)	(0.106)	(0.077)	(0.081)
Opinion: 0.75		0.478***	0.489***	0.573***	0.391***	0.606***	0.353***
		(0.051)	(0.078)	(0.112)	(0.085)	(0.069)	(0.069)
Observations	42,300	42,300	44,000	19,800	20,800	42,300	42,300

Notes: This table reports coefficients from conditional logit estimates by decision set (participant X block X period).

Robust standard errors, clustered at the individual level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table D6: Conditional Logit estimates of targeting a given node - Heterogeneity analysis, part B

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Engineer YES	Engineer NO	Raven Low	Raven High	Random Low	Random High	Mistakes Low	Mistakes High
Best Response	1.100***	1.278***	1.042***	1.302***	1.421***	0.790***	1.320***	0.940***
	(0.048)	(0.060)	(0.052)	(0.054)	(0.047)	(0.052)	(0.048)	(0.059)
Opponent's	-0.279***	-0.344***	-0.330***	-0.277***	-0.269***	-0.368***	-0.258***	-0.370***
target	(0.052)	(0.065)	(0.056)	(0.059)	(0.052)	(0.067)	(0.053)	(0.064)
Center	1.216***	1.250***	1.227***	1.234***	1.361***	1.050***	1.252***	1.197***
	(0.054)	(0.060)	(0.057)	(0.057)	(0.053)	(0.060)	(0.050)	(0.068)
Opinion: 0.25	-0.244***	-0.263***	-0.347***	-0.154**	-0.284***	-0.203**	-0.237***	-0.279***
	(0.076)	(0.081)	(0.080)	(0.077)	(0.075)	(0.082)	(0.070)	(0.091)
Opinion: 0.75	0.568***	0.364***	0.522***	0.451***	0.535***	0.437***	0.378***	0.639***
	(0.070)	(0.079)	(0.070)	(0.079)	(0.065)	(0.087)	(0.066)	(0.084)
Observations	49,400	35,200	41,200	43,400	54,200	30,400	53,400	31,200

Notes: This table reports coefficients from conditional logit estimates by decision set (participant X block X period). Robust standard errors, clustered at the individual level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table D7: Conditional Logit estimates of targeting a given node, conditional on the initial opinion distribution

	(1) All blocks	(2) Heterog.	(3) Heterog.
	Coeffs.	Coeffs.	Odds ratios
Best response	1.176***	1.173***	3.232***
	(0.038)	(0.044)	(0.143)
Opponent's target	-0.302***	-0.460***	0.631***
	(0.040)	(0.051)	(0.032)
Central node	1.229***	1.153***	3.169***
	(0.040)	(0.050)	(0.158)
Node's opinion 0.25: 1 *vs.* 1	-0.251	-0.241	0.786
	(0.173)	(0.166)	(0.131)
Node's opinion 0.25: 2 *vs.* 2	-0.284**	-0.284**	0.753**
	(0.117)	(0.114)	(0.086)
Node's opinion 0.25: 1 *vs.* 2	-0.306**	-0.282**	0.754**
	(0.120)	(0.118)	(0.089)
Node's opinion 0.25: 1 *vs.* 3	-0.215	-0.201	0.818
	(0.175)	(0.171)	(0.140)
Node's opinion 0.25: 2 *vs.* 1	-0.127	-0.114	0.892
	(0.103)	(0.100)	(0.089)
Node's opinion 0.25: 3 *vs.* 1	-0.581***	-0.565***	0.568***
	(0.132)	(0.129)	(0.073)
Node's opinion 0.75: 1 *vs.* 1	0.779***	0.759***	2.136***
	(0.123)	(0.119)	(0.253)
Node's opinion 0.75: 2 *vs.* 2	0.190*	0.187*	1.206*
	(0.109)	(0.104)	(0.126)
Node's opinion 0.75: 2 *vs.* 1	0.716***	0.699***	2.012***
	(0.100)	(0.097)	(0.194)
Node's opinion 0.75: 3 *vs.* 1	0.434***	0.421***	1.523***
	(0.144)	(0.138)	(0.210)
Node's opinion 0.75: 1 *vs.* 2	0.461***	0.459***	1.583***
	(0.089)	(0.087)	(0.137)
Node's opinion 0.75: 1 *vs.* 3	0.391***	0.373***	1.452***
	(0.133)	(0.132)	(0.191)
Observations	84,600	42,300	42,300

Notes: This table reports results from conditional Logit regressions by decision set (participant X block X period). Coefficients are reported in columns (1) and (2), and odds ratios in column (3). Robust standard errors, clustered at the individual level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table D8: Conditional Logit estimates of targeting a given node, conditional on the initial opinion distribution, for high- and low-Raven participants

	High-Raven			Low-Raven		
	(1)	(2)	(3)	(4)	(5)	(6)
	All blocks	Heterog.	Heterog.	All blocks	Heterog.	Heterog.
Best response	1.306***	1.311***	3.710***	1.043***	1.029***	2.798***
	(0.054)	(0.061)	(0.227)	(0.052)	(0.063)	(0.175)
Opponent's target	-0.277***	-0.386***	0.680***	-0.334***	-0.554***	0.575***
	(0.059)	(0.078)	(0.053)	(0.056)	(0.067)	(0.038)
Central node	1.239***	1.120***	3.066***	1.230***	1.205***	3.337***
	(0.056)	(0.070)	(0.215)	(0.058)	(0.071)	(0.237)
Node's opinion 0.25: 1 *vs.* 1	-0.171	-0.150	0.860	-0.369	-0.373	0.688
	(0.224)	(0.212)	(0.183)	(0.277)	(0.271)	(0.187)
Node's opinion 0.25: 2 *vs.* 2	-0.318**	-0.323**	0.724**	-0.253	-0.240	0.787
	(0.160)	(0.152)	(0.110)	(0.179)	(0.178)	(0.140)
Node's opinion 0.25: 1 *vs.* 2	0.029	0.046	1.047	-0.663***	-0.655***	0.519***
	(0.156)	(0.151)	(0.158)	(0.170)	(0.168)	(0.087)
Node's opinion 0.25: 1 *vs.* 3	-0.356	-0.341	0.711	-0.072	-0.056	0.945
	(0.242)	(0.239)	(0.170)	(0.238)	(0.236)	(0.223)
Node's opinion 0.25: 2 *vs.* 1	-0.018	-0.014	0.986	-0.232	-0.214	0.807
	(0.145)	(0.139)	(0.137)	(0.143)	(0.141)	(0.114)
Node's opinion 0.25: 3 *vs.* 1	-0.674***	-0.662***	0.516***	-0.493***	-0.473***	0.623***
	(0.189)	(0.187)	(0.096)	(0.182)	(0.178)	(0.111)
Node's opinion 0.75: 1 *vs.* 1	0.954***	0.922***	2.514***	0.565***	0.558***	1.747***
	(0.176)	(0.170)	(0.426)	(0.169)	(0.162)	(0.283)
Node's opinion 0.75: 2 *vs.* 2	-0.080	-0.082	0.921	0.488***	0.492***	1.636***
	(0.154)	(0.145)	(0.134)	(0.154)	(0.151)	(0.247)
Node's opinion 0.75: 2 *vs.* 1	0.782***	0.744***	2.105***	0.665***	0.667***	1.949***
	(0.157)	(0.153)	(0.322)	(0.126)	(0.122)	(0.238)
Node's opinion 0.75: 3 *vs.* 1	0.287	0.266	1.304	0.550***	0.548***	1.730***
	(0.204)	(0.196)	(0.256)	(0.199)	(0.190)	(0.329)
Node's opinion 0.75: 1 *vs.* 2	0.499***	0.492***	1.636***	0.427***	0.422***	1.524***
	(0.131)	(0.127)	(0.207)	(0.121)	(0.120)	(0.183)
Node's opinion 0.75: 1 *vs.* 3	0.246	0.231	1.260	0.510***	0.499**	1.648**
	(0.169)	(0.166)	(0.209)	(0.195)	(0.196)	(0.324)
Observations	43,400	21,700	21,700	41,200	20,600	20,600

Notes: This table reports results from conditional Logit regressions by decision set (participant X block X period). Participants are assigned to the high- or low-Raven characteristics based on the average Raven score. Coefficients are reported in columns (1) and (2), and odds ratios in column (3). Robust standard errors, clustered at the individual level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.