IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17315

Competing for Influence in Networks through Strategic Targeting

Margherita Comola
Agnieszka Rusinowska
Marie Claire Villeval

SEPTEMBER 2024

DISCUSSION PAPER SERIES

IZA DP No. 17315

Competing for Influence in Networks through Strategic Targeting

Margherita Comola
RITM, Paris School of Economics and IZA

Agnieszka Rusinowska
CES, CNRS - University Paris 1 Panthéon-Sorbonne and Paris School of Economics

Marie Claire Villeval
CNRS, Université Lumiére Lyon 2 and IZA

SEPTEMBER 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Competing for Influence in Networks through Strategic Targeting*

We experimentally investigate how players with opposing views compete for influence through strategic targeting in networks. We varied the network structure, the relative influence of the opponent, and the heterogeneity of the nodes' initial opinions. Although most players adopted a best-response strategy based on their relative influence, we also observed behaviors deviating from this strategy, such as the tendency to target central nodes and avoid nodes targeted by the opponent. Targeting is also affected by affinity and opposition biases, the strength of which depends on the distribution of initial opinions.

JEL Classification: C91, D85, D91

Keywords: network, influence, targeting, competition, laboratory experiment

Corresponding author:
Marie Claire Villeval
Univ Lyon
CNRS
GATE UMR 5824
F-69130 Ecully
France
E-mail: villeval@gate.cnrs.fr

* We are grateful 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 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, maximize one's impact, or secure strategic advantages is a fundamental challenge in many contexts, made even more complex when competing with rivals. For example, brands compete to convince key opinion leaders to endorse their products and disseminate information throughout their network. Politicians compete for endorsement from influential figures within a community to sway the opinion of groups of voters. Companies compete to form alliances with key partners who can provide them with a competitive edge for technology adoption or market access. In the realm of information warfare, organizations compete to influence individuals who can propagate partisan news on social networks and shape public opinion. In cybersecurity, attacking and defending critical nodes in a network can decide the success of an operation.

However, despite its empirical relevance and while there is 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] much less is known about how competition between influencers affects target selection in networks. Namely, in which contexts agents tend to target the same nodes as a competitor, when they flee away, and whether this is advantageous?

In economics, the theoretical literature on network targeting was pioneered by Ballester et al. (2006) that introduced a non-cooperative game in which players choose how much effort to exert. In this linear-quadratic framework, the Nash equilibrium strategy is proportional to the players' Bonacich centrality,[2] and the key player is identified using the intercentrality measure.[3] This model and its extensions have inspired theoretical investigations (*e.g.*, König et al., 2014; de Marti and Zenou,

[1] 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).

[2] This measure captures the influence of a player within the network based on both their direct connections and the connections of their neighbors.

[3] This identifies the player whose removal from the network would most significantly reduce overall efficiency or influence.

2015; Bloch and Shabayek, 2023), as well as empirical studies concerning a variety of social and economic 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), and finance (Battiston et al., 2012; Cabrales et al., 2016; Demange, 2018). However, one common feature of these empirical studies is that they examine cases where a single external actor has a monopoly on targeting.

In contrast, we focus on situations where multiple strategic agents compete in targeting. This competitive setting has received little attention. Grabisch et al. (2018) have described its theoretical properties in a model where two strategic agents with opposing opinions engage in a constant-sum game, targeting non-strategic agents who update their opinions according to the DeGroot model (DeGroot, 1974).[4] No empirical study has yet provided behavioral insights into multi-agent targeting. Building on Grabisch et al. (2018), our research aims to fill this gap in the context of information diffusion in networks.

We tested this Targeting Game in a laboratory experiment by forming exogenous social networks of five nodes. For each network, we introduced two strategic players with opposing opinions aiming to steer the average opinion of the nodes towards their own. To achieve this, each strategic player had to target one node in the network to form a link, thereby influencing that node's opinion and, indirectly, the opinions of its neighbors. Because our primary interest lies in the targeting decisions rather than the opinion-updating process, the nodes were represented by automated players. To facilitate decision-making, the opponent was also represented by an automated player, and participants were informed of the opponent's target choice before making their own. Once both strategic players selected their target nodes, information spread throughout the network, and opinions were updated using the DeGroot process.[5] This setup created a constant-sum game, where the two strategic

[4]See Rusinowska and Taalaibekova (2019) for an extension with three competing agents and de Vos et al. (2024) for a related model with finite-horizon opinion dynamics.

[5]Various methods of information processing can be found in the literature on opinion dynamics. These include Bayesian models, where agents use Bayes' rule to assess the state of the world (Banerjee, 1992; Bikhchandani et al., 1992; Gale and Kariv, 2003), non-Bayesian models initiated by DeGroot (1974), where agents imitate others by considering a weighted average of their neighbors' past opinions or actions (DeMarzo et al., 2003; Golub and Jackson, 2010, 2012), and models that

agents competed by choosing their targets.

The best-response strategy in this game depends on the relative influence of the strategic agents. High-influence agents should target central nodes, because they are within reach of the agent's influence and greatly impact other nodes. Low-influence agents should instead target peripheral nodes because they have less connections and are more influenceable. To explore this empirically, we manipulated the relative degree of influence of the two strategic agents. In the Increasing Influence treatment, the two strategic players started with the same degree of influence but the opponent's influence increased exogenously over the periods, while the player's influence remained constant throughout the experiment.

In addition to the opponent's degree of influence, the optimal targeting strategy depends in our setting on the opponent's target choice (a central *vs.* a peripheral node) and the network structure (either a line, a kite, a star, or a butterfly). Our design allows us to explore participants' choices across all scenarios. In total, participants played eight blocks of five periods each: within a given 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. Thus, both types of opponent's target can be observed across all degrees of influence.

Finally, we manipulated the nodes' original opinions across blocks between and within subjects. In half of the blocks, all nodes started with a neutral opinion (*i.e.*, they were indifferent between the two strategic players), while in the other half, they could be randomly assigned three different opinions. While theoretically irrelevant, we anticipated that players might still base their targeting decisions on these initial opinions. In our setting of DeGroot updating, as long as the two strategic players are assumed stubborn (*i.e.*, uninfluenced by the opinion of others), the initial opinion of nodes has no impact on the distribution of final opinions.

To ensure the robustness of our findings, we implemented two control treatments between subjects. In the Decreasing Influence treatment, the opponent's degree of

combine both approaches (Jadbabaie et al., 2012; Chandrasekhar et al., 2020).

influence decreased across rounds within a block instead of increasing. In the Layout treatment, we altered the visual display of the networks on the players' screens to prevent the central node in the network from appearing in the center of the screen.

Our design allows us to assess the extent to which players in a simplified laboratory setting adopt best-response strategies and to identify the behavioral factors that influence their choices when they deviate from these optimal strategies. The multi-agent setting uniquely positions us to explore these factors, particularly the tendency to target the same node as the competitor or, conversely, to flee away from it, even when such choices are not optimal. Furthermore, the preference for targeting the center of a network 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.[6]

Our study yields three key findings. First, the best response was selected approximately 70% of the time, showing an adjustment of choices based on the opponent's level of influence and the position in the network of their selected target. This result replicates in the Decreasing Influence and Layout control treatments, indicating that the rate of best responses was unaffected by the direction in which the competitor's influence evolves, or the visual presentation of the networks. However, the likelihood of best-response behavior was higher among participants with high mathematical and cognitive abilities, when networks had a clear focal point and multiple best responses, and in later blocks of the experiment, certainly due to learning.

The second and third main findings reveal behavioral regularities that deviate from equilibrium strategies. After controlling for the best responses, participants exhibited a residual tendency to target central nodes rather than peripheral ones and to avoid nodes also targeted by their opponents, suggesting a desire to differentiate themselves from the competitor. We also uncovered affinity and opposition biases. Participants showed a strong preference for targeting nodes with initial opinions closer to their own, and a (moderate) aversion to nodes with more distant opinions,

[6] In a DeGroot diffusion setting with one single targeting agent who is stubborn, the targeting choice of this agent does not affect the final distribution of nodes' opinions but only the speed of convergence to it.

despite these initial opinions having no impact on best-response strategies. The strength of these biases varies with the distribution of initial opinions. 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 being represented by one node only.

These findings highlight the complexity of strategic behavior in networks when influencers have to compete for influence. The tendencies to target the center and flee away from the competitor suggest that players may prioritize visibility and differentiation, even when this deviates from best-response strategies. The fact that the affinity and opposition biases are strongest when there is a minority or majority of close or distant opinions suggests that players are sensitive to the social composition of a network. Understanding these dynamics can inform the design of interventions aimed at promoting cooperation, reducing polarization, or enhancing the effectiveness of influence campaigns in various domains.

The remainder of this article is organized as follows. Section 2 presents a brief review of the related literature. Section 3 introduces our experimental design and procedures, while Section 4 develops our results. Section 5 discusses these findings and concludes.

2 Related literature

In addition to the articles mentioned in the previous section, our research connects to three strands of network literature: studies on targeting, studies on competition in targeting, and experiments on networks.

Our study primarily relates to the vast literature on diffusion (Jackson and Yariv, 2011; Lamberson, 2016), learning (Golub and Sadler, 2016), and contagion (Cabrales et al., 2016) in networks. In this area, several theoretical studies aim at identifying optimal targets to diffuse information. For example, in Galeotti and Goyal (2009), optimal influence strategies depend on the nature of the interaction (word-of-mouth communication *vs.* close collaboration), which leads to targeting either low or high connectivity individuals. In Chatterjee and Dutta (2016), a firm outside a network chooses an implant to propagate its product. A firm that produces a good quality product places its implant at a node that maximizes the decay centrality, while a firm producing a bad quality product places it at nodes with the highest number of connections. Focusing on optimal contracting, Belhaj and Deroïan (2019) show that a principal benefits from contracting with only a subset of the network, avoiding central agents when the interaction intensity is high (see also Belhaj et al., 2023). Tsakas (2017) models a planner's optimal targeting to maximize the diffusion of action in a society with agents who imitate others' successful past behavior. Studying how the position of the first person to receive information determines its diffusion in the network, Banerjee et al. (2013) model a word-of-mouth diffusion with an application to microfinance loan data. They define diffusion centrality, which measures how extensively information spreads from a node. They show that understanding the nature of transmission is key for identifying optimal injection nodes.[7]

Our study distinguishes itself by focusing on target competition and adopting an empirical perspective, an area that few studies address.[8] Among the exceptions, Bimpikis et al. (2016) model firms' targeted advertising in a competitive setting, 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 Robertson, 2021). In these

[7]There are also studies analyzing 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., 2023).

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

zero-sum games, two competing players allocate limited resources across multiple battlefields. Players face a trade-off: concentrate resources on a few battlefields to outperform the opponent, or spread resources more widely, risking defeat in key battles. While we share with Colonel Blotto games an interest in the strategic decision of where to exert influence under competitive pressure, our focus is, unlike in these games, on social networks with interconnected nodes. Our contribution lies in studying whether players can identify a single entry point that will maximize the dissemination of their opinion within the network, relative to their opponent's.

Adopting an experimental approach allows us to study whether competing individuals optimally respond to the network structure, as well as their opponent's relative power and targeting choices. If deviations from best responses occur, it helps us identify the behavioral patterns that lead to these deviations. Therefore, our study also contributes to the relatively scarce experimental literature on social networks (see surveys by Breza, 2016; Choi et al., 2016).

Most experimental studies focus on whether individuals play the equilibrium in network games but do not consider targeting.[9] A few experimental studies investigated how individuals form links with others, depending on the decision timing (Charroin, 2023), the network structure (He and Zou, 2024), or the presence of homophilous preferences (*e.g.*, Currarini et al., 2009; Charroin et al., 2022). However, in these studies, there is no competition between players when they create a link and not necessarily an intention to influence others. Additionally, experiments in psychology have shown that some individuals are more frequently targeted due to their personality, position within the network, and initial opinions.[10] While our ex-

[9] 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.

[10] 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, 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.

periment innovates by its approach, it also builds on the previous findings to explore the behavioral determinants of competitive targeting in networks.

3 Experimental design and procedures

3.1 Design

The Targeting Game. As in the theoretical model of Grabisch et al. (2018) that is summarized in online Appendix A, two strategic agents, called *Player* and *Opponent* respectively, 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 $[0, 1]$.[11] 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 he or she is connected to.

The two strategic agents (*Player* and *Opponent*) have fixed and opposite opinions on the matter of interest: the Player has opinion 1, and the Opponent has opinion 0. Their objective 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 reach this objective, they each create a link with one node of their choice in the network. 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 Player's degree of influence is fixed and standardized to 1, while the Opponent's degree of influence (denoted λ) varies. Here, $\lambda = X$ means that the Opponent's opinion counts X times more than that of a standard node or the Player. For example, if $\lambda = 1$, the Opponent's importance is equal to that of any other node or the Player; if $\lambda = 2$, the Opponent's opinion counts twice as much as that of a standard node or the Player.

[11] 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.

Once the Player and the Opponent have created a link with a node, the updating mechanism of nodes is a weighted average of opinions and converges (*i.e.*, produces a stable opinion vector) after a given number of periods. 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 own target). This is because we are interested in studying the best responses.

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]

Variations. In what follows we describe our main treatment, named Increasing Influence treatment. 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 exogenously varied the Opponent's degree of influence λ from 1 to 5 in a fixed ascending order over the five periods.

Between blocks, we exogenously varied the network configuration, the Opponent's targeting choice between center and periphery, and the nodes' initial opinions, 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 (see Figure C1 and the timeline

[12] Using computerized agents avoided prosocial considerations and strategic uncertainty that could affect Players' decisions. It facilitated identifying whether a participant played the best response.

[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, 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.

illustrated in Figure C2 in online Appendix C). Each network configuration appeared in two blocks, one block in which the Opponent targeted the center of the network and another block in which it targeted its periphery. The order of blocks was randomized at the individual level.

We also manipulated, within and between subjects, the nodes' initial opinions. All participants faced four blocks where all nodes' initial opinion was equal to 0.5 (*"homogeneous blocks"*) and four other blocks where the program randomly assigned different initial opinions to the nodes (0.25, 0.5, and 0.75) (*"heterogeneous blocks"*). In the heterogeneous blocks, there was an independent random draw for each node (and each subject), but the program ensured that each value was represented at least once in the network; thus, the number of nodes with any given opinion ranged from 1 to 3. The distribution of initial opinions was node- and block-specific and initial opinions were fixed within a block. Each player faced in random order four homogeneous blocks (two where the Opponent targeted the center and two where it targeted the periphery) and four heterogeneous blocks (two where the Opponent targeted the center and two where it targeted the periphery). Each network structure appeared in one homogeneous block and one heterogeneous block.

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 C3 in online Appendix C). Once the participant made a decision, a black avatar representing the 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 strategic players' targeting decisions. The value of each node's opinion and the color of the nodes changed continuously until convergence on the screen.

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 for that period.

Control treatments. Two features of the previously described design could influence behavior in the experiment. First, the Opponent's degree of influence always increased across periods within a block. 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 control treatments.

Unlike in the previously described *Increasing Influence treatment*, 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. The rest of the design was the same as in the Increasing Influence treatment. 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*, we altered the standard graphical layout by reshuffling the nodes' positions so that the central node was no longer depicted in the center of the graph (see the right column of Figure C1 in online Appendix C). This variation was only implemented for the Increasing Influence treatment. This is the only difference with this treatment. Comparing behavior between these two treatments helps determine whether a bias towards centrality is induced by the experimental design itself or by the inherent tendency of the participants.

Elicitation of cognitive abilities. We anticipated that cognitive abilities might influence participants' ability to best respond in the Targeting Game. Thus, we measured these abilities in several ways. In part 1, before the main game, 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 the winner of a €10 prize being the one closest to two-thirds of the session's average. The equilibrium is 0 and lower numbers indicate higher cognitive abilities. In part 3, after the main game, participants completed six Raven matrices in six minutes, earning ECU50 (€0.45) for each correct answer. The number of correct answers served as an additional measure of cognitive abilities.

3.2 Best responses

The theoretical predictions of the Targeting Game, as implemented in our experiment, in terms of best response are summarized in Figure 1.

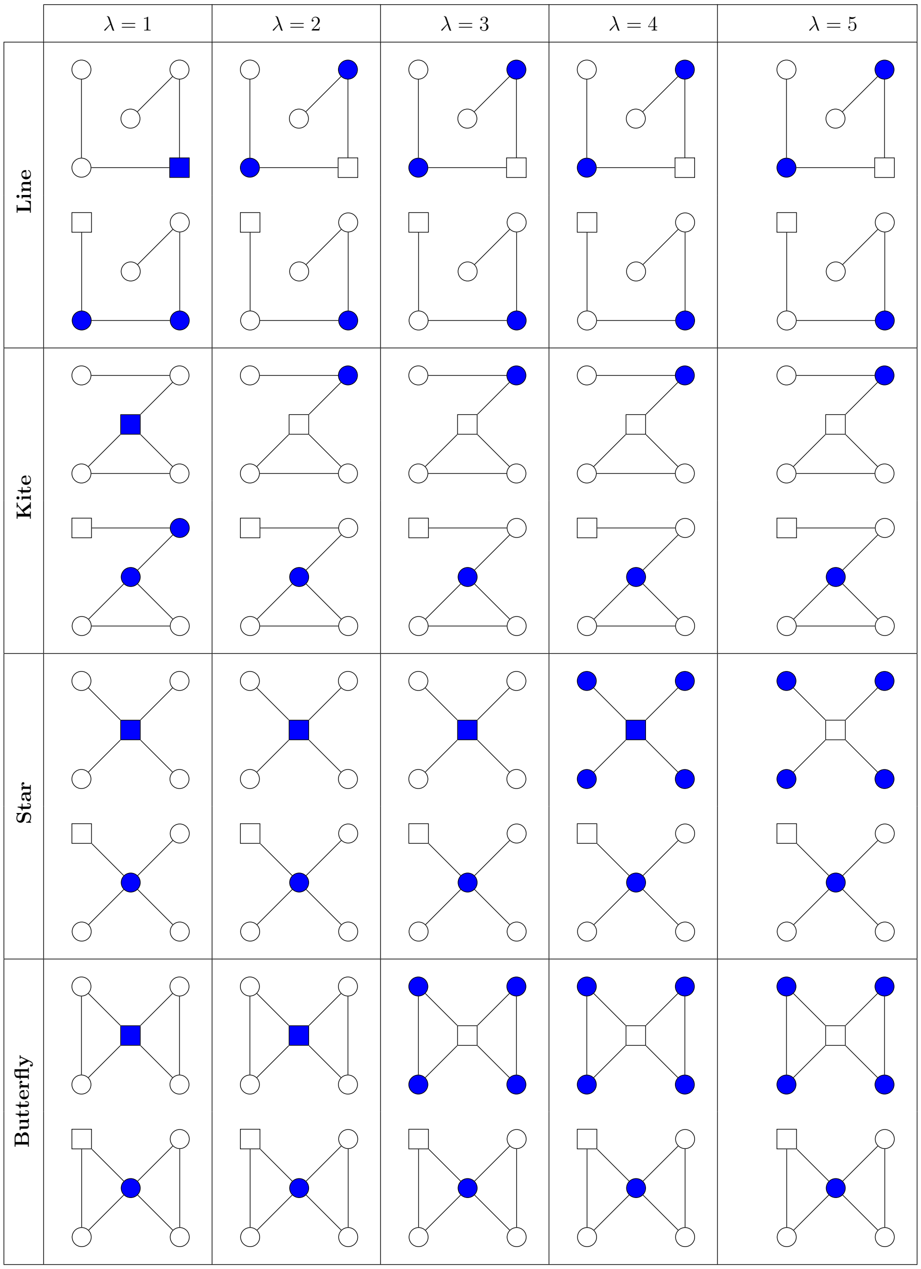
Each row in Figure 1 represents one of the 8 blocks, and each column represents one of the 5 periods. There are two rows for each network structure: one where the Opponent targets the central node, and one where it targets the periphery. Note that for any of our network structures, the main centrality measures assign the same node as the center.[14] The initial opinion of nodes is not reported in the figure because it does not affect the distribution of final opinions (and therefore the strategy profile). The columns in Figure 1 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 (which may consist of one or multiple nodes) by a circle.

The Player's best response is uniquely identified by the combination of three factors: the network structure, the Opponent's degree of influence λ and the Opponent's choice.[15] The best-response strategy is the same in the three treatments. As explained in online Appendix A, best-response targets are characterized by intermediacy centrality and influenceability (Grabisch et al., 2018), which account for the strategic nature of the two-player game. It 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

[14]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.

[15]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 1: Best responses in the Targeting Game



Note: The Opponent's target is depicted by a square and the Player's best response by a blue circle.

the center is optimal only while the Opponent's relative influence is low; the Player should switch away from the center when the Opponent's relative influence becomes sufficiently high. This implies that when λ becomes large enough, the Player should target a peripheral, *i.e.*, more influenceable, node, because it is the only one they can effectively influence (see the theoretical model in online Appendix A). However, the switch point 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 variability in the switching point provides a key source of identification in our design.

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.[16] Participants were primarily students from local engineering, business, and medical schools that we recruited via HRoot (Bock et al., 2014).[17] The experiment was programmed using Java.

Upon arrival at the laboratory, participants randomly drew a tag from a bag to assign them 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 (see instructions in online Appendix B.1). An experimenter read these instructions aloud. The instructions for the Targeting Game included several illustrations to help participants understand the weights of the nodes in the network, the Opponent's degree of influence, and the process of

[16]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.

[17]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.

opinion revision.

Before starting part 2, we checked the participants' understanding of the instructions using a quiz (see online Appendix B.1). Then, to familiarize them with the environment, participants had 6 minutes to practice the interface. During this practice time, they could observe a tree network, different from the network structures used in the experiment. 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 ("If this trial was paid, you would earn xx ECU").

The instructions for part 3 with the Raven matrices were displayed on the screen. After part 3, participants filled out a questionnaire with socio-demographic questions and queries about their social media usage, risk attitudes, and social preferences.

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 best responses. Next, we explore the determinants of choosing specific nodes in the network to identify behavioral regularities.

4.1 Best responses

Descriptive statistics. Table 1 summarizes the percentage of targeting choices that correspond to the best response in each treatment.

Participants chose the best response (or one of them, when multiple) 70% of the time in the Increasing Influence treatment. The relatively high frequency of best responses suggests that participants understood the game well. Pairwise comparisons with the Increasing Influence treatment show no significant differences in the rate of best responses compared to the Decreasing Influence treatment (69%; $p = 0.204$),

Table 1: *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
Opponent's 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
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. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

and the Layout treatment (68%; $p = 0.279$) (two-tailed Mann-Whitney rank-sum tests). The Layout and Decreasing Influence treatments differ significantly from the Increasing Influence treatment only when the network is a line ($p = 0.022$ and $p = 0.002$ respectively). This shows a negligible overall influence of these treatments on players' decisions. The rate of best responses differs significantly between the Decreasing and the Increasing Influence treatments also when considering each value of the λ parameter (at the 1% level when $\lambda > 1$). This is likely driven by the different points in time at which the participants face the need to switch their target in order to best respond. Given this limited number of differences, we pool all data in our econometric analysis and control for the treatment.

Table 1 also shows interesting differences in the best response rate across conditions. In particular, this rate is higher when the Opponent targets a peripheral node rather than the center of the network (76% *vs.* 65% in the Increasing Influence treatment; Wilcoxon signed-rank test, $p < 0.001$; no significant differences across treatments).

Table 1 also reveals a disparity of the best response rate 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, in the Increasing Influence treatment), that is, in settings where the network structure has a focal point and there are multiple best responses. Finally, the rate of best response differs according to whether the nodes have homogeneous or heterogeneous initial opinions (72% *vs.* 69 % in the Increasing Influence treatment; Wilcoxon signed-rank test, $p = 0.012$; no significant differences across treatments).

Regression analysis. We now explore the determinants of whether the participants played the best-response strategy, controlling for the characteristics of players and games. By pooling the data from the three treatments, we obtained a sample of 16,920 unique decision sets (423 participants X 8 blocks X 5 periods) where participants were called to target one 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 takes the value one if the targeting choice of player i for decision set d is a best response, and zero otherwise. All specifications include two dummy variables: one for the Decreasing Influence treatment and one for the Layout treatment, with the Increasing Influence treatment as the reference category. A dummy variable indicates whether the nodes' initial opinions were heterogeneous (the reference category is the homogeneous scenario where all initial opinions are equal to 0.5). X_i represents a vector of individual characteristics. In particular, three variables control for the participant's mathematical and cognitive abilities: a dummy variable indicating whether the player attends the Central engineering School in Lyon (a very selective school with high requirements in mathematics), the score in the Raven matrix task (ranging from 0 to 6), and the number chosen in the Beauty Contest game. We anticipated that higher cognitive abilities would facilitate strategic reasoning. Three additional individual characteristics are included: risk attitude, usage of social media (a dummy taking value one if the participant declares a moderate use, *i.e.*, below one hour per day, which proxies for experience in creating links in networks), and pro-sociality (re-scaled, with 1 unit representing 10 points donation). These regressions also control for the block number (ranging from 1 to 8), which can indicate learning in the game. The error term $\epsilon_{i,d}$ is clustered at the level of the participant. Table (2) reports the estimates from these regressions.[18]

We considered three specifications, depending on the fixed effects included. In column (1) we consider the specification of Equation (1), which includes no dummies for networks and Opponent choice and no fixed effects. In column (2), we added the Opponent's choice (dummy equal to one if the Opponent targeted the center, and 0 otherwise), dummies for each network structure (with the line network as the omitted category), and fixed effects for periods (from 1 to 5). In column (3), we replaced the Opponent's choice and the network dummies with a set of fixed effects representing the combination of network configuration, the Opponent's targeting choice (center *vs.* periphery), and λ (40 effects).[19] Since these three factors uniquely identify the best response in our game, this is our preferred specification for identifying the impact of randomly allocated experimental conditions.

Table (2) indicates no treatment differences at standard significance levels. It shows that participants were significantly less likely to best respond when the initial opinions of nodes were heterogeneous, which contrasts with the theoretical predictions. We explore this finding in a later subsection. Table (2) also provides evidence that mathematical and cognitive abilities are significant factors in the choice of the best response. These abilities, proxied by being a student at the Central engineering School, solving a higher number of Raven matrices, and reporting a lower number in the Beauty Contest game, increase the probability of best response play. As for the other individual attributes, we notice that a more prosocial attitude is associated with a lower likelihood of playing the best response (perhaps due to less strategic thinking), whereas risk attitude and usage of social media

[18] We also estimated a Probit model with similar specifications and independent variables as the linear probability model. The results are qualitatively similar and, therefore, are omitted but can be provided upon request.

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

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)	(0.018)	(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)
Opponent choice: Center	-	-0.108***	-
		(0.009)	
Network: Kite	-	0.003	-
		(0.011)	
Network: Star	-	0.134***	-
		(0.011)	
Network: Butterfly	-	0.100***	-
		(0.011)	
Constant	0.635***	0.606***	0.647***
	(0.037)	(0.040)	(0.043)
F.E. period	NO	YES	YES
F.E. network * Opponent's target * 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.

have no significant effect. Finally, results regarding the block number suggest some learning occurred during the game progression.

Results from column (2) indicate that identifying the best response is easier when the Opponent targets the periphery and when the network structure has a clear focal point and several best responses (star and butterfly networks), as already visible in Table 1. All the above results still 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 choose the best response in the Target Game, with this behavior being even more prevalent among those with higher mathematical and cognitive abilities.*

4.2 Targeting decisions

We now explore which characteristics of the nodes and the game influence target choices, aiming to identify behavioral tendencies beyond the best response.

Behavioral regularities. To identify behavioral regularities, we investigate the motivations behind choosing a specific node in a network. 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 the value of one 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. Therefore, we estimated a conditional logit model that can take 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 possible nodes in the decision set d, ensuring the probabilities sum up to one. 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 three specifications of this regression model. All 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.[20] Columns (2) and (3) also include the node's initial opinion (0.25 and 0.75, with 0.5 as the omitted category). Finally, in column (3), the sample is restricted to blocks with heterogeneous initial opinions of nodes.

Table 3: Probability of targeting a given node

Dependent variable: Choice of a node	(1) All blocks	(2) All blocks	(3) Heterogeneous blocks
Best response	1.159***	1.173***	1.166***
	(0.038)	(0.038)	(0.044)
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***
		(0.056)	(0.054)
Node's initial opinion: 0.75	-	0.488***	0.478***
		(0.052)	(0.051)
Observations	84,600	84,600	42,300

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

Table 3 shows, unsurprisingly, that a node is significantly more likely to be selected when it corresponds to the best response, confirming that participants have understood the game. However, even after controlling for the best response, there remains a residual tendency to target the center of the network rather than its periphery and to avoid targeting a node selected by the Opponent. This analysis yields our second result:

[20]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.

Result 2: *Controlling for best responses, there is a residual tendency to target the center of the network and to differentiate from the Opponent's target choice.*

Heterogeneity analysis. To check the robustness of these findings, we re-estimated the model specification reported in column (2) of Table 3, this time splitting the sample based on various individual and game characteristics (see Table D2 in online Appendix D).[21] We split the sample as follows: blocks with homogeneous initial opinions of nodes *vs.* blocks with heterogeneous opinions (columns 1 and 2); Increasing Influence treatment only (column 3); Decreasing Influence treatment only (column 4); Layout treatment only (column 5); earlier blocks *vs.* later blocks (columns 6 and 7); students at the Central engineering School *vs.* other students (columns 8 and 9); participants with above-average score in solving Raven matrices *vs.* those with equal or below-average scores (columns 10 and 11). The results in Table D2 are remarkably consistent with those from Table 3. This indicates that the overall findings are not driven by any particular subgroup of participants or specific game conditions.

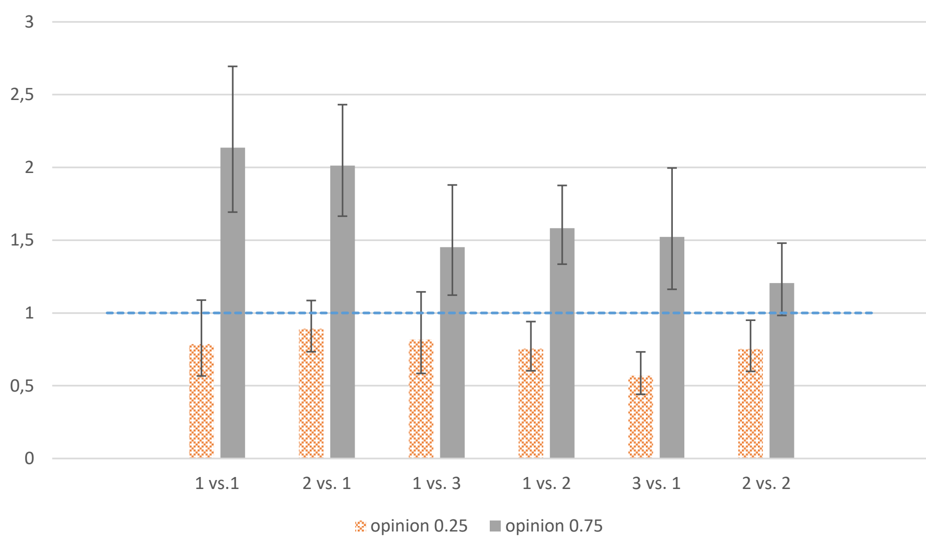
Affinity and opposition biases. Another intriguing behavioral pattern revealed by Table 3 concerns the influence of the nodes' initial opinions on targeting decisions, although, theoretically, these should have no impact. Specifically, participants were significantly more likely to target nodes with an initial opinion more similar to theirs (0.75), and less likely to target nodes with a more distant initial opinion (0.25), compared to a node with a neutral initial opinion (0.50).

However, the specifications in Table (3) overlook the possibility that the effect of the nodes' initial opinions may vary depending on the overall distribution of opinions among the five nodes. To recall, in each heterogeneous block, we randomly allocated the three initial opinion values (0.25, 0.5, and 0.75) to the nodes, with the constraint that each value had to be represented at least once. This process generated six possible distributions, which we denote by the number of nodes with 0.25 and 0.75 opinions, respectively. For example, in our terminology, '1 *vs.* 3' refers to a distribution where one node has a value of 0.25, three nodes have a value of 0.75, and the remaining node (implicitly assumed) has a value of 0.5.

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

In Table D3 in online Appendix D, we depart from the analysis in Table 3 by differentiating the impact of nodes' initial opinions (0.25 or 0.75) based on the overall distribution of opinions in the network. Column 1 reports conditional logit coefficients for the whole sample, while in column 2 we restrict to blocks with heterogeneous opinions of nodes. Column 3 reports the odds ratios from the regression of column 2, which eases the interpretation in what follows. We have plotted the odd ratios from column 3 (heterogeneous blocks only) in Figure 2. The omitted category (nodes with an initial opinion of 0.5) is depicted by a horizontal line at value 1, which serves as a reference for visually comparing with the implicit targeting probability of nodes with a neutral opinion. The dotted orange bars represent the odd ratios for nodes with an initial opinion of 0.25 while the dark grey bars represent those with an initial opinion of 0.75. We also report in black error bars for the 95% confidence intervals: when an error bar crosses the horizontal line, it suggests that the implicit targeting probability is not statistically distinguishable (at the 95% confidence level) from that associated with 'neutral' nodes with an initial opinion of 0.5.

Figure 2: Impact of the distribution of initial opinions in the network on the selected node



Notes: This figure reports the odd ratios from the conditional logit estimates reported in Table D3 in online Appendix D. The error bars show the 95% confidence intervals. Reading: for example, "2 *vs.* 1" reads as two nodes with initial opinion 0.25 and one node with opinion 0.75. To recall, the player's opinion is always equal to 1.

In Figure 2, overall, the dark gray bars indicate a strong preference among participants for targeting nodes with an initial opinion closer to their own, a tendency we refer to as 'affinity bias'. Conversely, the light bars reveal a (more moderate) aversion to targeting nodes with more distant opinions, which we call 'opposition bias'.

The relative magnitude of these biases varies depending on the opinion distribution within the network. Two situations are particularly noteworthy. First, in a balanced setting where both opinions are represented by multiple nodes (2 *vs.* 2 nodes, on the far right of the figure), the combined biases are at their weakest. In this setting, there is no significant affinity bias and only a moderate opposition bias. Second, conversely, these biases were more pronounced when the closer opinion was in a vulnerable position, that is, represented by one node only. In particular, when only one node shares a close opinion with the participant and the more distant opinion represents the majority (3 *vs.* 1), both affinity bias and the highest level of opposition bias are observed. On the other hand, when only one node shares a close opinion and the more distant opinion does not form a majority (1 *vs.* 1; 2 *vs.* 1, on the far left of the figure), affinity bias is strongest, while opposition bias is not significant. This analysis leads to our last result:

Result 3: *There is a strong preference for targeting nodes with opinions closer to one's own and a moderate aversion to targeting nodes with more distant opinions. 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.*

Efficiency. How costly are these behavioral biases in terms of efficiency? The online Appendix E provides a detailed analysis of efficiency measured by participants' payoffs. Recall that the best responses always paid 600 ECU. This analysis shows that deviations and biases were costly. In particular, participants' average earnings were lowest in the Heterogeneous condition when the best response was to target a node whose opinion was most distant from that of the player (mean = 572.2) and the highest when the best response was to target a node whose opinion was the closest to that of the player (mean = 580.8); the difference between the two averages is significant (Wilcoxon test, $p < 0.001$) and both are significantly below 600 ECU (t-test, $p < 0.001$) (see Figure E1). The analysis also shows that when players did not best respond, certain strategies led to lower payoffs than others (for example, targeting the central node paid only 520.1 on average, while targeting a peripheral node that was not targeted by the Opponent paid 540.4, both averages being significantly lower than 600 ECU (t-tests, $p < 0.001$) (see Figure E2)).

5 Discussion and conclusion

In this study, we investigated experimentally how competing influencers choose which node to target to maximize their impact on average public opinion in a network. We found that participants best responded to their opponent's targeting choices approximately 70 % of the time, taking into account their relative influence. The best response rate was significantly higher among participants with higher mathematical and cognitive abilities, and when the task was made easier, that is, when networks had a clear focal point and several possible best responses, such as star and kite networks. We also found evidence of learning over time. We identified several behavioral regularities that deviated from theoretical predictions. Controlling for the best response, we found a residual tendency to target the center of the network rather than peripheral nodes and avoid targeting a node that was already selected by the opponent, suggesting a differentiation strategy.

One of the most intriguing findings in our experiment is the evidence of affinity and opposition biases in targeting decisions. Players were more likely to target nodes with initial opinions more similar to their own and less likely to target nodes with a more distant opinion, even though the final average opinion in the network (and thus players' payoff) does not depend on the initial opinion of the targeted node. The strength of affinity and opposition biases depends on the distribution of initial opinions. These biases were less pronounced in a balanced setting where the opinions of the two strategic players were represented by multiple nodes, and they were highest when the opinion closer to the player's was represented by one node only.

This could be explained by at least three reasons. One possible reason is homophily and group identity. The literature on homophily has shown that individuals are more likely to form links with others who share similar traits or beliefs, either due to a taste for similarity or because they believe it is strategically advantageous (*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). Research on minimal group identity has revealed that individuals who identify with a particular group tend to favor members of this group and may exhibit hostility towards those outside of it (*e.g.*, Tajfel and Turner, 1986; Akerlof and

Kranton, 2000; Benjamin et al., 2010; Chen and Chen, 2011; Chen et al., 2014). This could explain why participants were more likely to target nodes with similar opinions, perhaps perceived as in-groups, and avoid those with different opinions, perceived as out-groups.

A second possible reason is the transposition of behavior from real-world social networks into the laboratory setting. When individuals form connections on social media, they typically invite or accept 'friend' requests from people with whom they perceive some degree of proximity. Participants may unconsciously replicate this behavior, targeting nodes that feel closer to them, as they would when choosing friends on social networks. We discard this interpretation, however, since the variable capturing the use of social media was never significant in our regressions.

A third possible reason relates to the perceived influenceability of the targeted node. In real life, an influencer might view individuals with similar opinions as easier targets because they are more likely to be receptive to their message, although such targets may yield low returns in terms of shifting the overall public opinion. Conversely, individuals with differing opinions might be seen as more challenging to persuade, although they could have a more significant impact in shifting aggregate public perception. How decision-makers navigate this trade-off remains an open empirical question that lies beyond the scope of our current investigation.

We acknowledge the potential limitations of our study regarding external validity, and thus we refrain from drawing overly broad conclusions about its real-world implications. Nevertheless, our results suggest that if people care about the proximity of the individuals they try to influence, even when these factors are irrelevant to their outcomes, balanced networks, characterized by a diversity of opinions, allow for a more efficient targeting strategy. The composition of networks in terms of initial opinions impacts the strength of the affinity and opposition biases, which may have implications for understanding influence dynamics in real-world networks. These insights may help us understand how information spreads, how consensus is reached, and how conflicts evolve. For example, in social media, users might prioritize engaging with like-minded individuals while avoiding confrontation with opposing views, thus contributing to echo chambers. In organizational or political settings, leaders might focus on influencing key individuals that align with their views, while minimizing conflict with opponents, which could affect coalition-building.

Our study represents a first step toward understanding targeting competition, a topic that has been overlooked in the literature despite its practical relevance. There are many possible extensions to our work. In particular, we used the simple DeGroot mechanism for modeling the updating of opinions. It would be valuable to explore other updating mechanisms that might, for example, factor in the relative weight of neighbors, or the distance to strategic agents, allowing for memory failures. Another avenue would be to explore targeting competition where strategic agents can remove a node from a network to disrupt the information flow, rather than creating a link to disseminate information. This is an area for further research.

References

Acemoglu, D., A. Ozdaglar, and J. Siderius (2021). Misinformation: Strategic Sharing, Homophily, and Endogenous Echo Chamber. *NBER Working Paper 28884*.

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). Diffusion of microfinance. *Science 341, DOI: 10.1126/science.1236498.*

Banerjee, A. V. (1992). A simple model of herd behavior. *Quarterly Journal of Economics 107(3)*, 797–817.

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.

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.

Benjamin, D. J., J. J. Choi, and A. J. Strickland (2010). Social identity and preferences. *American Economic Review 100*(4), 1913–1928.

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

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 two experiments. *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.

Charroin, L. (2023). Heterogeneity in sequential network formation games. *Journal of Behavioral and Experimental Economics 107*, 102075.

Charroin, L., B. Fortin, and M. C. Villeval (2022). Peer effects, self-selection and dishonesty. *Journal of Economic Behavior & Organization 200*, 618–637.

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.

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.

de Vos, W., P. Borm, and H. Hamers (2024). Influencing opinion networks: Optimization and games. *Dynamic Games and Applications 14*, 959–980.

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.

DeMarzo, P., D. Vayanos, and J. Zwiebel (2003). Persuasion bias, social influence, and unidimensional opinions. *Quarterly Journal of Economics 118*, 909–968.

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 (2023). Brokering votes with information spread via social networks. NBER Working Paper.

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.

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

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

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.

He, S. and X. Zou (2024). Public goods provision in a network formation game. *Journal of Economic Behavior & Organization 218*, 104–131.

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.

Jadbabaie, A., P. Molavi, A. Sandroni, and A. Tahbaz-Salehi (2012). Non-bayesian social learning. *Games and Economic Behavior 76*, 210–225.

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. Robertson (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.

Rusinowska, A. and A. Taalaibekova (2019). Opinion formation and targeting when persuaders have extreme and centrist opinions. *Journal of Mathematical Economics 84*, 9–27.

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.

Tajfel, H. and J. C. Turner (1986). The social identity theory of intergroup behavior. In S. Worchel and W. Austin (Eds.), *Psychology of Intergroup Relations*, pp. 7–24. Nelson Hall.

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[22] at time step t.

Following DeGroot (1974), Grabisch et al. (2018) associate to 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 $\lambda > 0$.[23] 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 at bringing the asymptotic average opinion in the network as close as possible to his own opinion (1 for a_1 and 0 for a_2).

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

[23]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

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

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 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[24] 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.[25]

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.[26] 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.

[24]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.*, $\prod_{k=1}^{K-1} w_{i_k, i_{k+1}}$.

[25]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.

[26]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 or not the same level of influence.

Grabisch et al. (2018) characterize equilibria in pure strategies under equal levels of influence ($\mu = \lambda$). A pair of strategies (i,i) is an equilibrium of the game $\mathcal{G}_\mu$ 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 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. 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.

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. Let us 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.

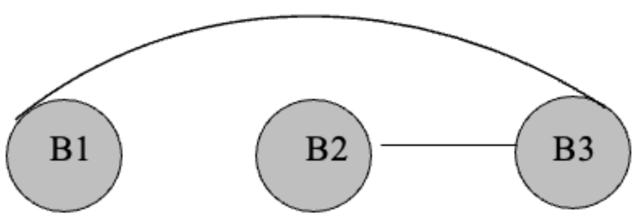
The weight of an opinion indicates the relative importance given by a participant B to the opinion of the participant with whom he or she has **a direct link**.

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).

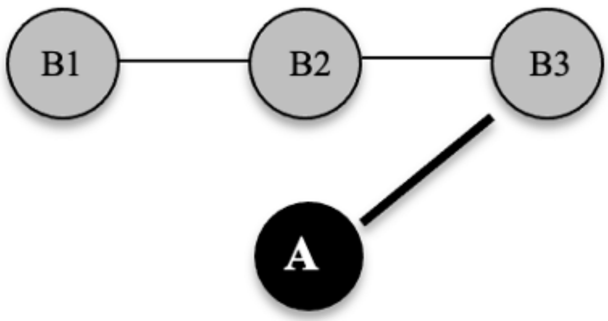
Degree of influence of participants A

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

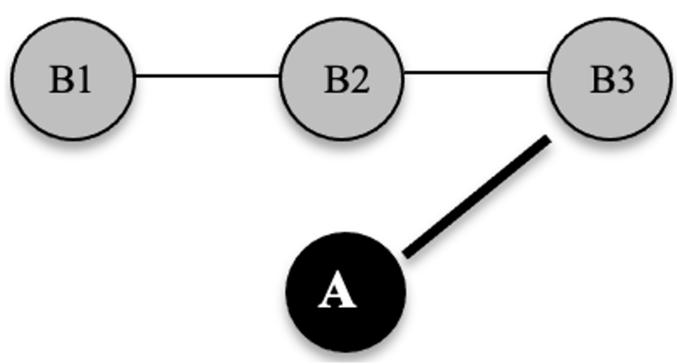
After having consulted the degree of influence of your Opponent, **you must choose one of the five participants B** by clicking on the grey circle representing him or her. You can choose the same participant as your Opponent.

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.

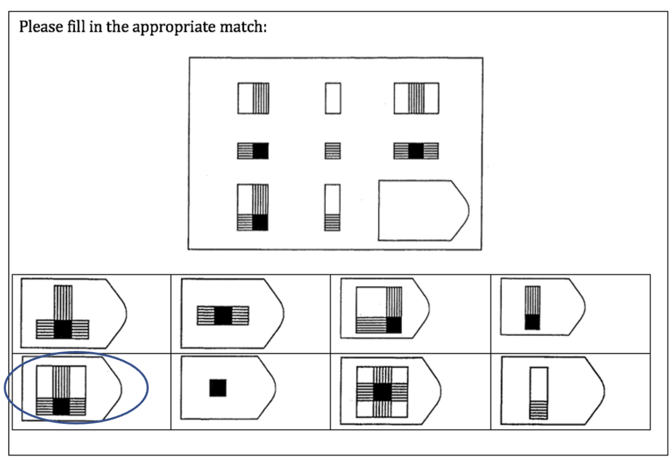
For each question, a series of figures will appear. Your task is to identify which figure logically follows the previous ones.

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.

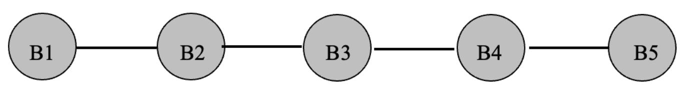
In the example below, the correct answer is circled.



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

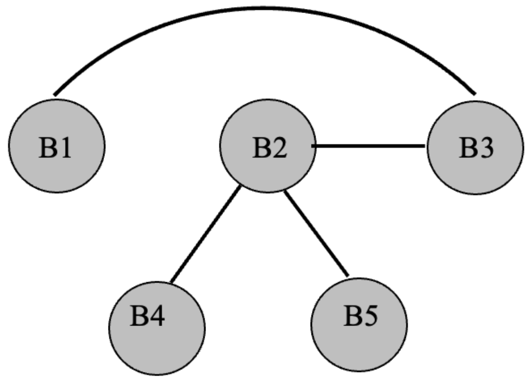
Please answer the following questions.

- 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:



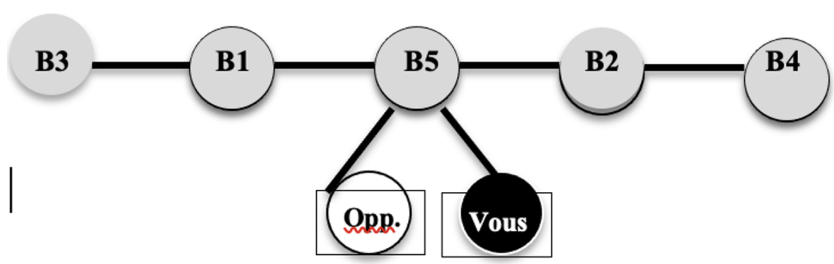
  - What is the weight of B2's opinion on B1? *(1)*
  - What is the weight of B2's opinion on B3?*(0.5)*
  - What is the weight of B4's opinion on B3? *(0.5)*
  - What is the weight of B4's opinion on B5? *(1)*

- 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 Ã©coles
  - 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 - ThÃšse
- 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

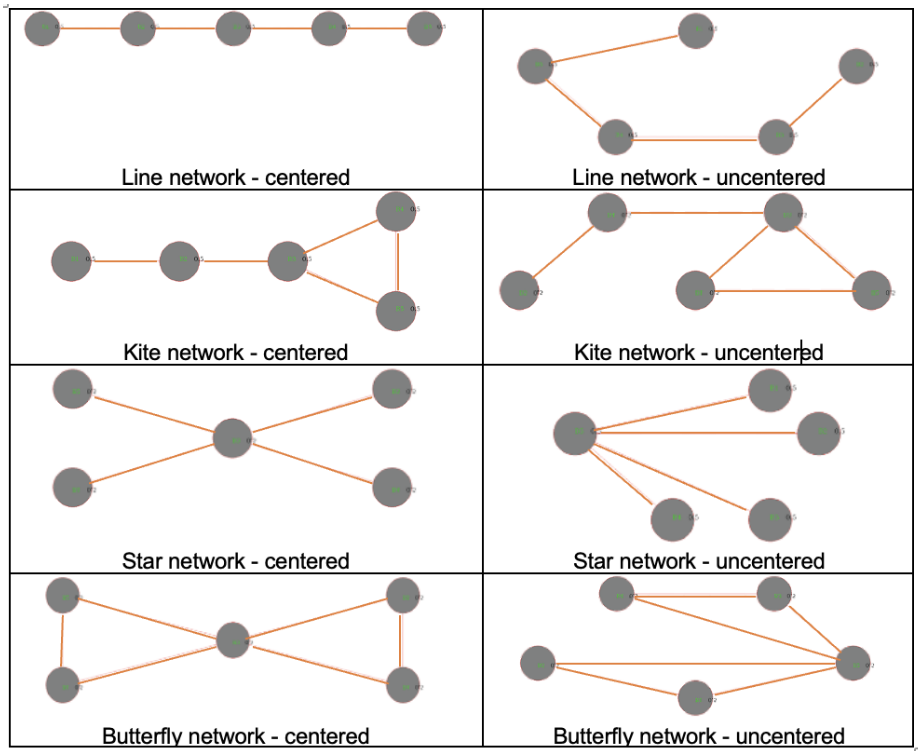


Figure C1: Networks structures in the experiment

Notes: The figure displays the four network structures used in the experiment as they appeared on the participants' screens. In the Increasing Influence and Decreasing Influence treatments, networks were displayed as shown in the left column. The Layout treatment displayed them as shown in the right column. The only difference in the Layout treatment was that the central node was not positioned in the center of the participants' screen.

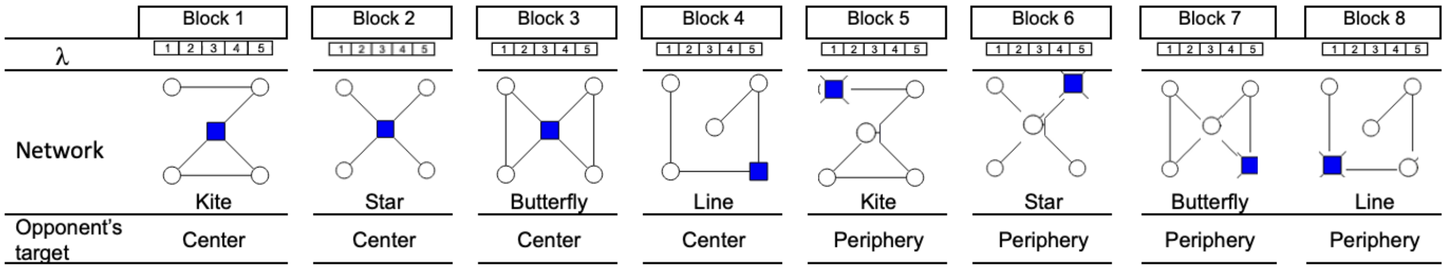


Figure C2: Timeline of the experiment

Notes: The figure illustrates the timeline of the experiment in the Increasing Influence treatment, with its 8 blocks. Within each block, there are 5 periods corresponding to an increasing degree of influence of the Opponent (λ). Each network structure appeared twice, once with the Opponent targeting the center, and once with the Opponent targeting a peripheral node. The Opponent's target is depicted by a blue square. In the experiment, the order between blocks was randomized at the individual level. The only difference with the Decreasing Influence treatment is that λ decreased over periods instead of increasing. We omit the initial opinions of the nodes for clarity of the figure.

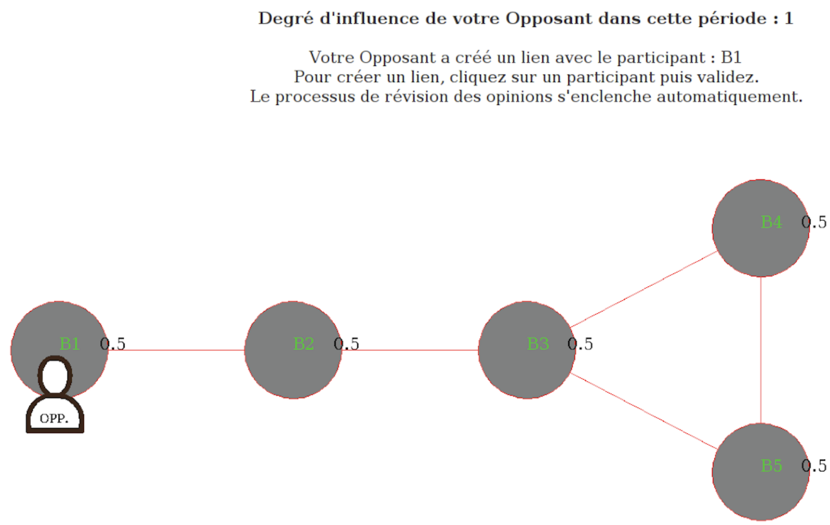


Figure C3: Example of a participant's screen

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.

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: Probability of targeting a given node - Heterogeneity analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Opinion Homog.	Opinion Heterog.	Treat. Increas.	Treat. Decreas.	Treat. Layout	Blocks: 1 to 4	Blocks: 5 to 8	Engineer YES	Engineer NO	Raven Low	Raven High
Best response	1.198***	1.166***	1.217***	1.181***	1.128***	1.034***	1.312***	1.100***	1.278***	1.042***	1.302***
	(0.049)	(0.044)	(0.055)	(0.082)	(0.076)	(0.050)	(0.046)	(0.048)	(0.060)	(0.052)	(0.054)
Opponent's	-0.141***	-0.463***	-0.135**	-0.749***	-0.256***	-0.286***	-0.340***	-0.279***	-0.344***	-0.330***	-0.277***
target	(0.049)	(0.051)	(0.055)	(0.079)	(0.081)	(0.049)	(0.061)	(0.052)	(0.065)	(0.056)	(0.059)
Center	1.288***	1.153***	1.291***	1.131***	1.166***	1.201***	1.279***	1.216***	1.250***	1.227***	1.234***
	(0.047)	(0.050)	(0.056)	(0.097)	(0.070)	(0.056)	(0.057)	(0.054)	(0.060)	(0.057)	(0.057)
Opinion: 0.25	-	-0.238***	-0.178**	-0.343***	-0.328***	-0.249***	-0.264***	-0.244***	-0.263***	-0.347***	-0.154**
		(0.054)	(0.079)	(0.115)	(0.106)	(0.077)	(0.081)	(0.076)	(0.081)	(0.080)	(0.077)
Opinion: 0.75	-	0.478***	0.489***	0.573***	0.391***	0.606***	0.353***	0.568***	0.364***	0.522***	0.451***
		(0.051)	(0.078)	(0.112)	(0.085)	(0.069)	(0.069)	(0.070)	(0.079)	(0.070)	(0.079)
Observations	42,300	42,300	44,000	19,800	20,800	42,300	42,300	49,400	35,200	41,200	43,400

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

Table D3: Probability of targeting a given node, conditional on the initial opinion distribution

Dep. variable:	(1) All blocks	(2) Heterog.	(3) Heterog.
	Coeffs.	Coeffs.	Odd 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.* 1	-0.284**	-0.284**	0.753**
	(0.117)	(0.114)	(0.086)
Node's opinion 0.25: 1 *vs.* 3	-0.306**	-0.282**	0.754**
	(0.120)	(0.118)	(0.089)
Node's opinion 0.25: 1 *vs.* 2	-0.215	-0.201	0.818
	(0.175)	(0.171)	(0.140)
Node's opinion 0.25: 3 *vs.* 1	-0.127	-0.114	0.892
	(0.103)	(0.100)	(0.089)
Node's opinion 0.25: 2 *vs.* 2	-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.* 1	0.190*	0.187*	1.206*
	(0.109)	(0.104)	(0.126)
Node's opinion 0.75: 1 *vs.* 3	0.716***	0.699***	2.012***
	(0.100)	(0.097)	(0.194)
Node's opinion 0.75: 1 *vs.* 2	0.434***	0.421***	1.523***
	(0.144)	(0.138)	(0.210)
Node's opinion 0.75: 3 *vs.* 1	0.461***	0.459***	1.583***
	(0.089)	(0.087)	(0.137)
Node's opinion 0.75: 2 *vs.* 2	0.391***	0.373***	1.452***
	(0.133)	(0.132)	(0.191)
Observations	84,600	42,300	42,300

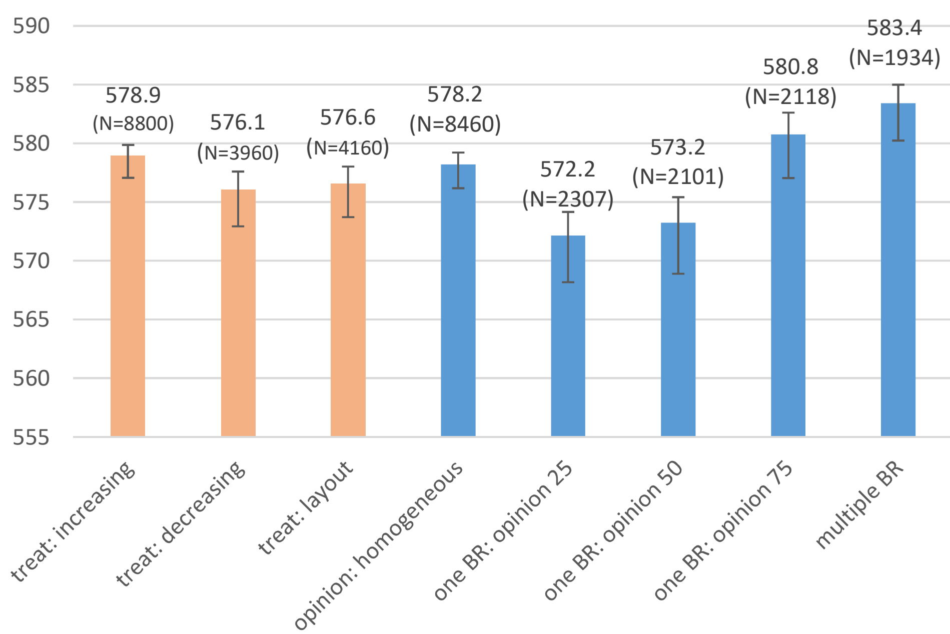
Notes: This table reports results from conditional Logit regressions by decision set (participant*block*round). Coefficients are reported in columns (1) and (2), and odd ratios in column (3). Robust standard errors, clustered at the individual level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

E Online Appendix: Efficiency analysis

In this section we analyze efficiency, as measured by the participants' earnings in the Targeting Game. Participants were paid proportionally to the maximum attainable payoff, with those who best responded always earning 600 ECU, serving as a natural efficiency metric. Figure E1 shows the average payoffs across all games and players by condition, plotted with 95% confidence interval bars. This figure reflects the efficiency of play based on the distance from the hypothetical average payoff of 600 ECU.

The results in FigureE1 align with earlier findings in Table 1, showing that payoffs are slightly higher in the Increasing Influence treatment than in the Decreasing Influence and Layout treatments, although the difference is not significant at standard levels.

Figure E1: Average payoffs across treatments and conditions



Notes: This figure reports the average payoffs, by condition, with 95% confidence interval bars. The unit of observation is the unique targeting choice (N=16,920). The first three bars correspond to the Increasing Influence, Decreasing Influence, and Layout treatments, respectively. The fourth bar is for the homogeneous opinion condition while the three following bars correspond to the heterogeneous opinion condition when the best response ("BR") is unique, depending on whether it corresponds to a node with opinion 0.25, 0.5, and 0.75, respectively. The far-right bar is for the heterogeneous opinion condition when several best responses coexist.

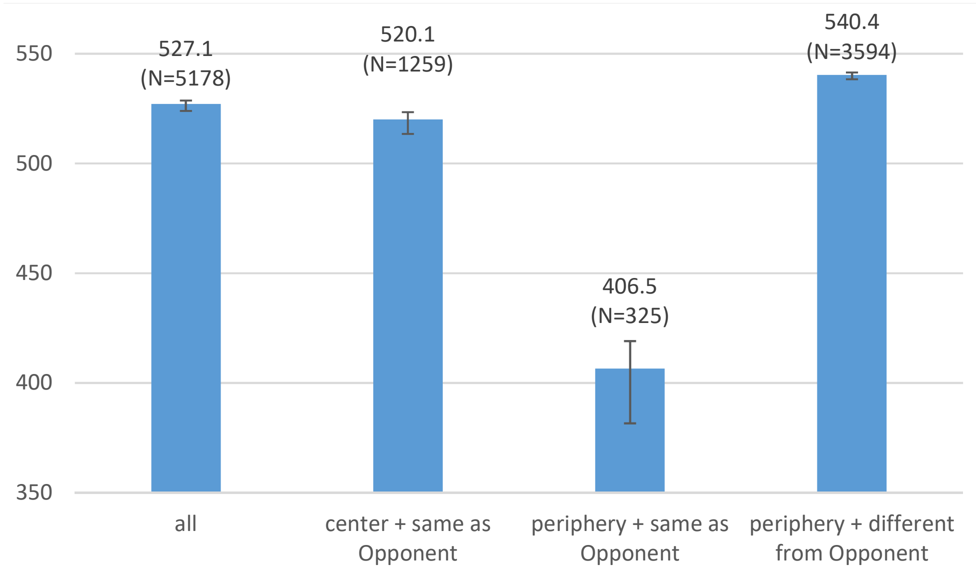
The analysis of payoffs conditional on the initial opinion of best-response nodes delivers interesting aggregate patterns. Since the initial opinions of nodes were randomly distributed and did not influence the best response, any differences in average payoffs reflect participants' behavioral biases in choosing which node to target. Average payoffs for the homogeneous opinion condition compare well with those of the Increasing Influence treatment. In the heterogeneous condition, we split the games based on the initial opinion of the best-response node(s). The bar labeled 'one BR: opinion 25' represents games where the initial opinion of the unique best-response node was 0.25, with bars labeled 'one BR: opinion 50' and 'one BR: opinion 75' are defined similarly. The last bar groups games where multiple nodes, possibly with different initial opinions, were a best response. The payoff for this residual category is mechanically higher because the latitude for errors is reduced.

In the Heterogeneous condition, participants' average payoffs were lowest when the best response was targeting a node with an opinion most distant from their own (mean= 572.2) and highest when the best response was targeting a node with an opinion closest to their own (mean= 580.8). The difference between these averages is significant (Mann-Whitney rank-sum test, $p < 0.001$), and both averages are significantly below 600 ECU (t-test, $p < 0.001$). In summary, average payoffs when the best-response node had an initial opinion closer to the participant's own opinion (0.75) were 8 points higher than in the 0.25 scenario, which gives a measure of the impact of the affinity and opposition biases.

Finally, in Figure E2 we plot the participants' payoffs when they did not best respond in total (N=5178), and in different situations. By focusing on the sub-sample of decisions that differ from the best responses, we can further explore the efficiency costs associated with each behavioral deviation identified in the individual-based analysis. We classify deviations into three types of 'wrong' choices: targeting the center like the Opponent, targeting a peripheral node like the Opponent, and targeting a peripheral node while the Opponent targets the center.[27] The results in FigureE2 show that most wrong choices occurred when participants targeted a different node than their Opponent, particularly when they targeted the periphery while the Opponent targeted the center, even when it would have been optimal to do otherwise. Targeting the central node when it is not the best response paid only 520.1 on average, while targeting periphery when Opponent targets center paid 540.4 on average, both averages being significantly lower than 600 ECU (t-test, $p < 0.001$). Although rare (N=325), instances where participants follow the Opponent in targeting a peripheral node result in the worst payoff outcomes (406.5).

[27]Note that, when the Opponent targets the periphery, targeting the center is always the best response. Thus, "Center + different from Opponent" is not a category in this figure.

Figure E2: Average payoffs when the selected target is not the best response



Notes: This figure shows the average payoffs when participants did not choose the best response, broken down by condition, with 95% confidence interval bars. The unit of observation is the unique targeting choice. The second and third bars represent cases where participants wrongly targeted the center or the periphery, respectively, mirroring the Opponent's choice. The far-right bar indicates the case where participants wrongly targeted the periphery while the Opponent targeted the center.