University of Cyprus
Department of Economics

Working Paper 01-2023

Communication and the Emergence of a Unidimensional World

Philippos Louis, Orestis Troumpounis and Nikolas Tsakas

Department of Economics, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus
www.ucy.ac.cy/econ

Communication and the emergence of a unidimensional world *

Philippos Louis[†], Orestis Troumpounis[‡], Nikolas Tsakas[§]

May 17, 2023

Abstract

While individuals hold, exchange, and update opinions over multiple issues, opinions are often correlated and a *unidimensional* spectrum is enough to summarize them. But when should one expect opinions to be unidimensional? And how important is the underlying structure of communication? Our experimental results: i) validate the crisp predictions by DeMarzo et al. (2003) when individuals update their opinions on a fixed network always trusting the same neighbors, ii) jointly with simulations indicate the prevalence of unidimensionality as an expected outcome even when communication is less structured with individuals' network possibly varying over time, and iii) highlight the importance of the communication structure in predicting whether individuals hold relatively moderate or extreme opinions.

Keywords: *opinion dynamics, information aggregation, persuasion bias, social networks.*
JEL codes: *D83, D85.*

*We are grateful to Dimitri Vayanos, Jeffrey Zwiebel, Dimitrios Xefteris, Elias Tsakas and seminar and workshop participants at the universities of East Anglia, Carlos III de Madrid, Cyprus, Lancaster, Loyola, Maastricht, Manchester, UNSW, Zurich. We acknowledge financial support of the Swiss National Science Foundation (SNSF 135135), the European Research Council (ERC Advanced Investigator Grant, ESEI-249433), the Spanish Ministry of Science and Innovation (FPI Grant BES-2012-054732) and Singapore University of Technology and Design (Project Code IDG31300110).

†Corresponding author: Department of Economics, University of Cyprus, PO Box 20537, 1678 Nicosia, Cyprus. E-mail: `louis.philippos@ucy.ac.cy`

‡Department of Economics and Management, University of Padova, Via del Santo 33, 35123 Padova, Italy & Department of Economics, Lancaster University, Bailrigg, Lancaster, LA1 4YX, United Kingdom.

§Department of Economics, University of Cyprus, PO Box 20537, 1678 Nicosia, Cyprus

1 Introduction

Social scientists tend to agree that the world is not flat; it is (often) unidimensional! The world we refer to is that of *opinions*, where although individuals hold opinions on a myriad of issues, spanning domains such as politics, the economy, or lifestyle, often a *unidimensional* spectrum can accurately summarize those. We encounter a good example in politics. Describing someone as a leftist or a rightist may provide enough information about their opinions on an array of political issues (e.g., redistribution, attitude towards the environment, abortion, immigration or gun possession). Indeed, individuals' opinions on different issues are often correlated, and the underlying ideology can explain legislators' and voters' behavior (Poole and Rosenthal, 1997; Lee et al., 2004; Ansolabehere et al., 2008). Ideological cleavages may further spillover to preferences over leisure activities, consumption or even art, as well as personal morality (see Kosinski et al., 2013, DellaPosta et al., 2015 and references therein, as well as Wilson and Haidt, 2014).[1]

The prevalence of a unidimensional world raises the question regarding its origin. DeMarzo, Vayanos, and Zwiebel (2003) provide a seminal model of opinion updating in a fixed network of communication and show how the repeated communication among boundedly rational individuals over multidimensional opinions can reduce disagreement on a single dimension. Despite the attractiveness of such an argument, there is no empirical evidence corroborating their theoretical findings: Can communication give rise to a unidimensional world? And what is the role of the communication structure in determining the shape of disagreement? In this paper, we precisely answer the above questions by testing and validating in a laboratory experiment the theoretical predictions of DeMarzo et al. (2003). Moreover, we provide and experimentally validate a conjecture on the generilizability of their results when communication networks are not fixed.

Figure 1 is useful in fixing ideas. It summarizes the opinions of five individuals across two issues: the practice of veganism and CrossFit training.[2] Panel *A* represents a multidimensional world, with no apparent correlation between issues. Panels *B* and *C* represent two different unidimensional worlds, where opinions on the two issues are strongly correlated. However, while we observe *unidimensionality* in both panels *B* and *C*, the shape of disagreement varies across the two, due to individuals' different *relative positions*. Individuals 1 and 5 have extreme positions in panel *B* on the opposite side of each other. While 1 remains an extremist in panel *C*, 5 has a moderate position.

DeMarzo et al. (2003) predict that repeated communication over a fixed network of communication among boundedly rational individuals makes opinions move from a multidimensional world (panel *A*) towards a unidimensional one (panels *B* and *C*). This requires

[1]Of course, it would be naive on our side to claim that opinions are always unidimensional and this is by no means the scope of the paper, see for example Page (2020) on instances where this is not the case.

[2]We pick two random issues to illustrate the concept of unidimensionality. Notice that in our setup opinions can be defined in a very broad sense including beliefs, judgements, attitudes and all such fundamental drivers of behaviour that are amenable to social influence or the advent of new information.

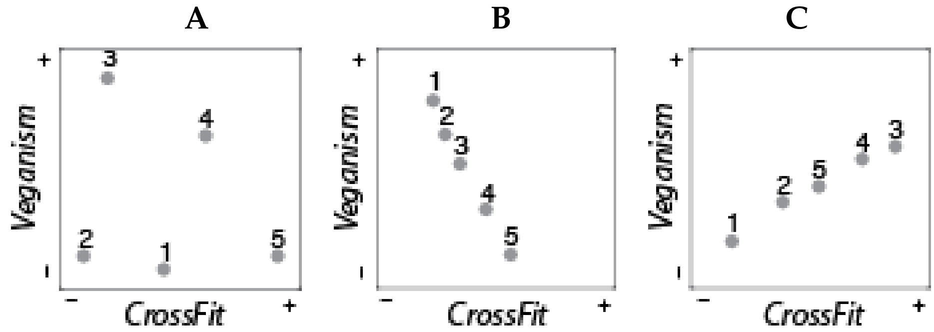


Figure 1: Each point represents an individual's opinions on the two issues: CrossFit and veganism. Opinions can vary from extremely negative to extremely positive. Panel A shows an example of uncorrelated opinions. Panels B and C show examples of unidimensional opinions.

individuals to update their opinions as in a DeGroot (1974) type model where they communicate their opinions on an array of issues over several rounds, and update their opinions by taking a weighted average of their own prior opinion and those of others. In such a process, opinions in the limit converge (DeGroot, 1974; Golub and Jackson, 2010).[3] But as DeMarzo et al. (2003) highlight, prior to opinion convergence, disagreement can be summarized by a unidimensional spectrum. DeMarzo et al. (2003) also provide crisp predictions that can differentiate between panels *B* and *C* depending on the characteristics of a fixed network of communication. Namely, the fixed structure of such a network determines individuals' opinions' relative positions.

Our experimental results not only support the predictions in DeMarzo et al. (2003) but also suggest that communication gives rise to unidimensionality in more instances than the (specific) ones known so far. In particular, our experiment validates our simulations that permit: a) communication channels to vary over time and hence the absence of a *fixed* network structure, and b) individuals to potentially assign different weights to one's opinions in different rounds. That is, in contrast to DeMarzo et al. (2003), individuals need not communicate in every round with the same individuals, nor assign the same weight to others' opinions in all rounds. Both generalizations seem relevant and realistic: individuals do come across different interlocutors over time, and even if they meet the same ones may not always trust them in the same manner. Our simulations show that even in such flexible communication environment, unidimensionality remains the prevalent outcome. At the same time, individuals' relative positions cannot be predicted any longer, unless one imposes further restrictions on the way communication structures vary over time. The simulations demonstrate that some predictability of relative positions can be recovered by limiting the variability of the communication environment to networks with a fixed distribution of neighbors or a constant underlying structure.

[3] A long literature focuses on the necessary and sufficient conditions for a society to reach consensus, as well as the speed at which this can be achieved. See for instance Golub and Jackson (2010, 2012); Tahbaz-Salehi and Jadbabaie (2008), or more recently Molavi et al. (2018) for an axiomatic approach. For a broad overview see Jackson (2008) Chapter 8.3 and references therein.

Our experimental design permits subjects to communicate their opinions over two issues (across ten rounds in groups of five). In two treatments, subjects were linked through a (different) fixed network (similar to Corazzini et al. 2012; Brandts et al. 2015). In a third treatment, subjects listened the opinions of different randomly picked subjects in each round. In line with the theoretical framework, our results support the emergence of a unidimensional world in all three treatments, and also the role of networks in determining individuals' relative positions. Nevertheless, we still find some divergence between theoretical predictions and experimental data. In particular, while unidimensional worlds emerge prominently in some groups, they fail to do so in others. To understand whether this divergence is due to individual behavior not conforming to the averaging model, we extend the simple homogeneous model to allow for heterogeneity. We use clustering analysis to classify subjects into different types. We then show through simulations that once possible heterogeneous behavior is taken into account, the averaging model can closely replicate our experimental results. Taken together, our results provide support for the emergence of unidimensionality through communication, but also highlight the critical role of heterogeneity in this process.

Our experimental design and results complement previous experiments where communication in a fixed network is over a single issue and the analysis focuses on the individual updating process showing that simple averaging models are performing well (see Corazzini et al. 2012; Brandts et al. 2015; Battiston and Stanca 2015; Chandrasekhar et al. 2020; Grimm and Mengel 2020). This is particularly so when the network is unknown to subjects (Grimm and Mengel, 2020), a feature we maintain in our laboratory setting. In terms of design, our work is the first to permit communication of opinions over more than one issue and therefore a focus on the emergence of a unidimensional world and its characteristics. In terms of results and subjects' behavior, although we also show that some kind of opinion averaging is a prevalent updating process, we also highlight the importance of permitting heterogeneity in the updating behavior.

Our work further links to different strands of the literature. In line with the DeGroot (1974) model of updating and the subsequent work by DeMarzo et al. (2003) individuals fail to take into account correlated information, the focus of vibrant ongoing research on correlation neglect and its consequences (e.g., Levy and Razin 2015, 2016, 2018, 2019) motivated by recent empirical (e.g., Ortoleva and Snowberg 2015) and experimental evidence (e.g., Enke and Zimmermann 2019). Turning attention to the unidimensionality result, McMurray (2014) shows how political competition can lead to a unidimensional policy space for parties. Spector (2000) shows how unidimensional beliefs can emerge in a model of sequential cheap-talk communication preceding a collective decision where individuals end up agreeing in all but one issue, while in our setting unidimensionality arises while disagreement is still present in all dimensions.

The rest of the paper is structured as follows. Section 2 presents the theoretical background and our results based on simulations. Section 3 presents the experimental setting and Section 4 our experimental results. Section 5 concludes.

2 Theoretical background

2.1 An Example

Before proceeding with the necessary notation and formal definitions, we provide a simple example that illustrates the relevant notions. Consider a network of five individuals where the channels of communication among agents over several rounds are represented by a directed network (Figure 2). The arrows represent the direction of information flow, so for example, the arrow that starts from agent 5 and points at agent 3 means that agent 3 listens to the opinion of agent 5.

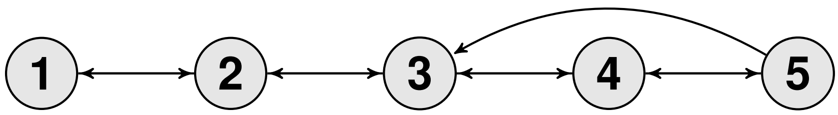


Figure 2: A network of five agents. Each node corresponds to an agent. An arrow pointing from one agent to another indicates that the latter listens to the opinion of the former.

Let agents have some initial opinions in round $t = 1$ over two issues (x, y) as represented at the left part of Figure 3. Based on the DeGroot (1974) weighted averaging model, let agents updated opinions be a weighted average of their own and all their neighbors' opinions in the previous round. Let also agents put equal weight on the opinions of themselves and each of their neighbors. For example, since agent 2 observes the opinions of agents 1 and 3, their updated opinion in each round and each issue will be a weighted average of the opinions of agents 2, 1, and 3, each with equal weight (1/3). Starting with the initial opinions in Figure 3 on the left, we depict all agents' opinions after two rounds of updating ($t = 2$ in the middle and $t = 3$ on the right).

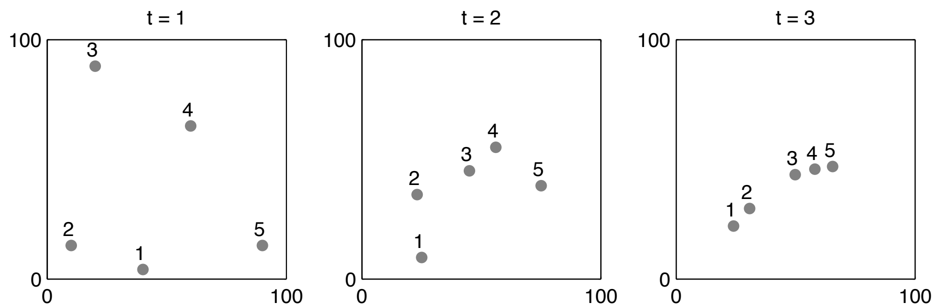


Figure 3: An example of opinion updating for agents in the network depicted in Figure 2. At $t = 1$, opinions are drawn randomly and independently for each agent and in each dimension. They are represented as points on a plane, with numbers indicating each agent's position in the network. After two rounds of updating, at $t = 3$, agents' opinions lie roughly along a line, with agent 1 in the lower left position, followed by others in increasing order.

Let us now focus on the properties of interest. First, notice that while at $t = 1$ opinions are scattered around the plane without a specific pattern, by $t = 3$ all opinions are aligned. This is precisely the notion of *unidimensionality* this paper deals with. Moreover, notice that the ordering along this line is with 1 and 5 at the two opposite extremes and 3 in the middle. As it turns out, these *relative positions* are not random and actually depend on the assumed updating process, the network structure, and the weights agents put to each other's opinion.

2.2 The model

Our theoretical framework pertains to the family of average–based updating processes introduced by DeGroot (1974), and builds upon DeMarzo et al. (2003). Consider a population of N agents, forming opinions on K different issues.[4] Agents communicate in discrete time rounds $t \in \{1, 2, \dots\}$ and update their opinions. Their initial opinions at time $t = 0$ are given exogenously. The opinion of agent i on issue k at time t is $s_{i,k}(t) \in \mathbb{R}$ and the $N \times 1$ column vector $\mathbf{s}_k(t)$ denotes the opinions of all agents on issue k at round t. We summarize all agents' opinions in all dimensions at time t by the $N \times K$ matrix $\mathbf{s}(t)$, where $\mathbf{s}(0)$ is the matrix of initial opinions.

Communication: at every round $t \in \{1, 2, \dots\}$, each agent i observes the opinions across all K issues in round $t - 1$ from a subset of the population. Communication may not be reciprocal, i.e., an agent may be observed by some agent they do not observe and vice versa. The collection of all those connections defines the communication network in round t. We restrict attention to strongly connected networks, which is necessary to allow the flow of information in the population.[5] In principle, the network may change over time. We say that a network is *fixed* if each agent observes the same set of agents in all rounds; otherwise, we say that the network is *varying*.

Updating: At each round $t \in \{1, 2, \dots\}$, each agent's opinions are a weighted average of their own opinions at round $t - 1$ and the rest of the opinions they observed. The matrix $\mathbf{T}(t) = \left(T_{i,j}(t)\right)$ is called a *listening matrix* and its typical element $T_{i,j}(t)$ represents the weight that agent i puts on agent j's opinion at round t. $T_{i,i}(t)$ is the weight i puts on their own opinion, which we assume to be strictly positive, while in general $T_{i,j}(t) \in [0, 1]$ and $\sum_{j=1}^{N} T_{i,j}(t) = 1$. Notice that if the network is varying, then by definition the listening matrix varies as well. However, the listening matrix may still vary even if the network remains fixed. We denote a finite and an infinite sequence of listening matrices respectively as follows: $\mathcal{T}^t = \{\mathbf{T}(\tau)\}_{\tau=1}^{t}$ and $\mathcal{T}^\infty = \{\mathbf{T}(\tau)\}_{\tau=1}^{+\infty}$.

The distinction between the network and the listening matrix reflects the two important

[4] In the current context, it is perhaps best to think of opinions as agents' estimates about the state of nature, as this is how we consider opinions in our experiment. Another alternative would be to consider them as preference parameters, like attitudes towards different choice alternatives. In any case, the internal consistency of the model of opinion dynamics is not affected by their exact meaning, as it is a purely mechanical process.

[5] A network is said to be strongly connected if there is a directed path from any agent i to any other agent j in the network.

ingredients of the opinion-updating process. The latter captures the structure of social interaction within the population: *who observes whom?* The listening matrix adds to that the behavioral elements of opinion updating: *how much weight one puts on their neighbors' opinions and how this changes at each point in time*. This distinction becomes particularly important in our experimental setup. While, in the lab, it is possible to shape the communication network, it is not possible to control the weight each subject puts to others.

We can now formalize the opinion updating process in its general form as follows:

$$\mathbf{s}(t+1) \;=\; \mathbf{T}(t+1)\cdot\mathbf{s}(t) \tag{1}$$

where (1) can be also written as:[6]

$$\mathbf{s}(t+1) \;=\; \prod_{\tau=1}^{t+1}\mathbf{T}(\tau)\cdot\mathbf{s}(0) \tag{2}$$

Unidimensional opinions: The aim of the paper is to study the possibility that opinions become unidimensional as a result of the described opinion formation process and to identify the role of the communication network in such a process. Opinions are said to be *unidimensional* when the points describing each agent's opinion on the k dimensions all fall on a straight line traversing $\mathbb{R}^k$ (as in Figure 3). To formalize this idea we introduce some notation related to *principal component analysis (PCA)*.

In this setup, we are particularly interested in the parameter $\beta^P(t) \in \left[\frac{1}{K}, 1\right]$, which is the *percentage of total variance of the multidimensional opinion profile that is explained by the first principal component* at time t.[7] If all variance can be explained by the first principal component, then opinions can be summarized by a single dimension, i.e., opinions are unidimensional. Formally,

Property 1. (Unidimensionality) An opinion formation process that can be described by (1) has the *unidimensionality* property when

$$\lim_{t\to\infty}\beta^P(t) = 1$$

The *unidimensionality* property formalizes the idea that communication can lead to a unidimensional world, i.e. that we can describe the differences of opinions using a single

[6]As matrix multiplication is in general non–commutative, it should be clarified that we consider *backwards matrix products*, i.e. $\prod_{\tau=1}^{t}\mathbf{T}(\tau) = \mathbf{T}_t\cdot\mathbf{T}_{t-1}\cdot\dots\cdot\mathbf{T}_1$.

[7]The value of $\beta^P(t)$ can be calculated as follows: Let $\hat{s}(t) = s(t) - \mathbf{1}\,\bar{s}(t)$, where $\bar{s}(t)$ is an 1 x K vector containing the mean opinion in each dimension at time t and $\mathbf{1}$ is a N x 1 vector of ones. Let $PC_n(t)$ be the eigenvector corresponding to the n-highest eigenvalue of the K x K covariance matrix of $\hat{s}(t)$. $PC_1(t)$ is the 1^{st} principal component of opinions at time t. Then, $s^P(t) = (\hat{s}(t)\cdot PC_1(t))^T$ is the projection of agents' opinions on this principal component. Now, to calculate $\beta^P(t)$ one needs to calculate the projection of $\hat{s}(t)$ on all principal components and take the covariance matrix of that. This is a diagonal matrix and $\beta^P(t)$ is the ratio of the first element of the diagonal over the sum of all diagonal entries.

line, like the left-right spectrum in politics (e.g., panels *B* and *C* of Figure 3). As evident from the graph, and the discussion so far, one could also state unidimensionality in terms of correlation of opinions across dimensions. Nevertheless, PCA allows us to state the remaining properties below without assuming that unidimensionality holds, as one can always project opinions on the first principal component. It also proves to be a valuable tool for the analysis of our experimental data for the same reason.

Relative positions: Notice that *unidimensionality* ensures that the relative position of each agent with respect to all others will be the same on all issues. This comes as a direct result of the linear relation between opinions. Thus, once *unidimensionality* is achieved relative positions on the different issues can be summarized by an agent's position on the line. In the absence of *unidimensionality*, one can describe relative positions through the opinions' projections on the first principal component.

We capture this via the property of *Position Determinacy*. More specifically, let $\mathbf{C}^{\kappa,\lambda}(t)$ be the $N \times N$ *opinion comparison matrix* at round t whose element $C^{\kappa,\lambda}_{i,j}(t)$ is equal to 1 whenever i's opinions' projection on the first principal component relative to j's is concordant to κ's opinions' projection of the first principal component relative to λ's at round t, and equal to 0 otherwise.[8] It follows from (2) that an opinion comparison matrix depends on the sequence of listening matrices and on the initial opinions: $\mathbf{C}^{\kappa,\lambda}(t) = C^{\kappa,\lambda}(\mathcal{T}^t, \mathbf{s}(\mathbf{0}))$. *Position Determinacy* implies that in the long run relative positions converge in a way that depends only on the sequence of listening matrices $\mathcal{T}^\infty$, and not on the initial opinions $\mathbf{s}(0)$. Formally:

Property 2. (Position Determinacy) An opinion formation process that can be described by (1) satisfies *Position Determinacy* if

$$\lim_{t\to\infty} \mathbf{C}^{\kappa,\lambda}(t) = C^{\kappa,\lambda}(\mathcal{T}^\infty)$$

where $\mathcal{T}^\infty$ is the sequence of listening matrices and (κ, λ) is an arbitrary reference pair.

To grasp the essence of the property more clearly, consider, for example, panels *B* and *C* of Figure 1. Comparing the opinions of individuals 3 and 5, one can say that while in Panel *B* individual 5 has more extreme opinions than individual 3 the opposite holds in Panel *C*. Picking individuals 1 and 3 as the reference pair, the projections of individuals on the first principal component would generate the following opinion comparison matrices:

[8] A formal definition would be as follows: Consider the following *relative comparison function*

$$c(x, y; \kappa, \lambda) = \begin{cases} 1, & if\ x > y\ \&\ \kappa > \lambda \\ & or\ x < y\ \&\ \kappa < \lambda \\ 0, & otherwise \end{cases}$$

The *opinion comparison matrix* $\mathbf{C}^{\kappa,\lambda}(t)$ has elements $\mathbf{C}^{\kappa,\lambda}_{i,j}(t) = c(s^1_i(t), s^1_j(t); s^1_\kappa(t), s^1_\lambda(t))$.

$$\hat{\mathbf{C}}^{1,3}_{Panel_B} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \hat{\mathbf{C}}^{1,3}_{Panel_C} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}$$

Focus on element (5,3) comparing the positions of 5 and 3 using as a reference 1 and 3. From Panel C, this element takes value 1 in $\hat{\mathbf{C}}^{1,3}_{Panel_C}$ since 5's opinions' projection on the first principal component relative to 3's is concordant to 1's relative to 3's. From Panel B in contrast, this element takes value 0 in $\hat{\mathbf{C}}^{1,3}_{Panel_B}$ since, 5's opinions' projection on the first principal component relative to 3's is not concordant to 1's relative to 3's. The same exact difference between the matrices is observed when comparing the positions of 4 and 5 across the two panels. All remaining elements are the same.

Relative positions can be summarized by other, perhaps simpler, measures. The advantage of the opinion comparison matrix for our study will come to light in the analysis of the experimental data. There it allows for a direct comparison of relative positions of pairs of agents in different treatments.

Unidimensional opinions in fixed networks: When in all rounds $t = 1, 2, \ldots$ (i) the network of communication is strongly connected, (ii) all agents put strictly positive weights to their own current opinion, and (iii) all positive weights are uniformly bounded away from zero from below, then convergence of opinions to consensus is guaranteed in the long–run (see Nedic and Ozdaglar, 2009; Nedic and Liu, 2014).

DeMarzo et al. (2003) study a particular class of such updating processes where agents communicate over a fixed network and may change over time the relative weight they put on their own opinion compared to the weights they put on others. This process is summarized by (2) with the relevant updating listening matrix defined as follows:

$$\mathbf{T(t)} = (1 - \lambda(t))\,\mathbf{I} + \lambda(t)\mathbf{T} \tag{3}$$

In the expression above, $\mathbf{I}$ is the identity matrix, $\mathbf{T}$ is a listening matrix that remains fixed for all t, and $\lambda(t) \in (0, 1]$ denotes the relative weight assigned to ones' own individual prior opinion compared to the listening matrix $\mathbf{T}$. It turns out that as long as agents do not become "too stubborn, too early"[9] the analysis can be concentrated on the properties of $\mathbf{T}$. Denoting by α_2 the second largest eigenvalue of this matrix the following result is obtained:[10]

Theorem 1 (restatement of Theorem 4, DeMarzo et al. (2003))**.** *Consider a generic listening matrix* $\mathbf{T}$ *with* $\alpha_2 \in \mathbb{R}$*. Then, the opinion process described by (3) satisfies Unidimensionality and Position Determinacy.*[11]

[9]The formal necessary condition for the results is that $\sum_{t=1}^{\infty} \lambda(t) = +\infty$.

[10]The eigenvalues are ranked according to their modulus, as they might be complex numbers. The modulus $||\alpha||$ of a complex number $\alpha = a + ib$ is $||a + ib|| = \sqrt{a^2 + b^2}$.

[11]In particular, agents' long run positions are determined by the values of the elements of the second column

Unidimensional opinions in varying networks: Theorem 1 makes sharp predictions on the rise of unidimensionality and agents' positions assuming that the listening matrix remains the same in all rounds. It remains an open question whether these properties of the process are robust to the network varying across rounds. Motivated by simulation results, we put forward the following conjectures:

1. *Unidimensionality* can still arise in the process described by (3) even when the listening matrix varies in each round.
2. Predictions about the long-run *relative positions* of the agents, when the listening matrix is varying, may be possible when adding structure to the sequence of possible listening matrices.

In what follows we present the results of four sets of simulations giving rise to the above conjectures that are also relevant for our experiment. Based on standard network formation models, these illustrate how unidimensionality arises frequently (Figure 4) and how imposing more structure on the network formation process allows for predictions concerning relative positions (Figure 5).

Each set of simulations consists of 1000 trials on networks of 50 agents who update their opinions over two issues according to (3) for up to 150 rounds.[12] In all four sets of simulations, the networks are generated randomly in each round, using different network generating processes, and are restricted to be strongly connected.[13] In the first two sets of simulations no further structure is imposed except for that dictated by the network formation process: the first set (SN) contained only Scale-Free networks, while the second set (SW) contained only Small-World networks.[14] In the other two sets of simulations, we imposed some further core structure across rounds. More specifically, in the third set (FDSF), similar to our experiment, we consider again randomly generated directed Scale-Free networks, but we keep the in-degree distribution fixed across all rounds and all trials.[15] In the fourth set (LSW), we consider (what we call) directed Linear Small-World networks. These are networks obtained in the same way as Small-World networks, with the underlying structure being an undirected linear network, rather than a circular one.[16] Upon construction of the

eigenvector of $\mathbf{T}$.

[12]We stopped the updating process if the range of opinions in at least one issue was lower than 5×10^{-12}. Initial opinions were drawn uniformly from $[0, 100]$.

[13]We have repeated the exercise using other network generating processes, as well as other parameter values. The results are qualitatively similar to the ones presented and they are available upon request.

[14]Scale-Free networks were constructed using the Barabàsi-Albert model (Barabási and Albert, 1999) with three new links added at every step. Small-World networks were constructed using the Watts-Strogatz mechanism (Watts and Strogatz, 1998), with each agent being initially connected with the three closest agents in each side and the destination of each directed link being rewired with probability 0.2.

[15]To form the initial in-degree distribution, we used the Barabàsi-Albert model. Subsequently, in each round, each agent would observe the opinions of a subset of the population with size equal to its in-degree. The subset was chosen uniformly at random among all subset of size equal to the agent's in-degree.

[16]Agents were again initially connected with their three closest neighbors on each side, except those at the edges of the line.

network and for all four sets, the weights of the listening matrix were randomly generated, with weight to one's own opinion being bounded below by 0.2.[17]

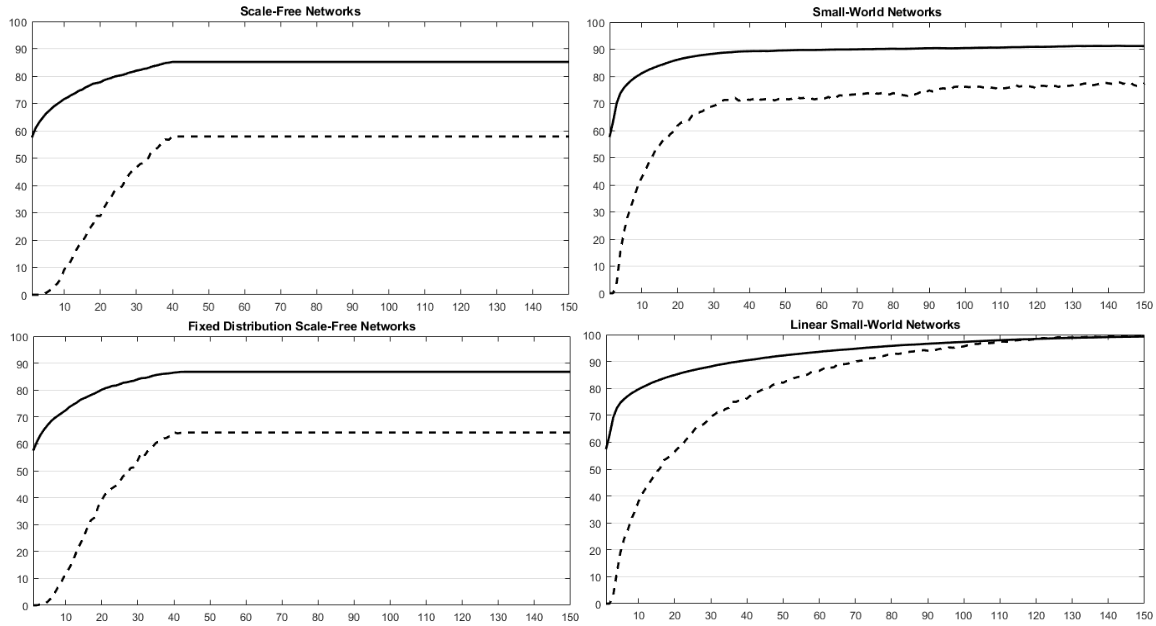


Figure 4: The horizontal axes measure the rounds, while the vertical axes depict percentages. Solid lines show the average $\beta^P(t)$ (in percentage terms) over the 1000 trials for each round t. Dashed lines show the percentage of trials in which $\beta^P(t)$ was larger than 85% in round t.

Regarding the unidimensionality property, Figure 4 illustrates the evolution of the percentage of the variance explained by the first principle component $\beta^P(t)$ across the 150 rounds of updating. The graphs show a large increase over rounds both of the average $\beta^P(t)$ –solid lines– and of the share of trials that reach $\beta^P(t) > 85\%$ –dashed lines– for all four set of simulations. Of course, neither all sets of simulations result to unidimensionality at the same speed, nor all set of simulations result to same percentages. However, the graphs illustrate that unidimensionality is to be expected as a frequent outcome of the assumed updating process, even when the listening structure varies.

Figure 5 instead focuses on agents' relative positions. The horizontal axis in the graphs represents the agents' index. In simulations SF and FDSF we index agents based on their in-degree in the first round, from smallest to largest (breaking ties randomly). In SF, as the in-degree varies randomly in each round, this order is arbitrary. On the other hand, in FDSF, since the in-degree remains constant across rounds, agents with a higher index have a higher in-degree throughout the process. In simulations SW and LSW the index corresponds to the agent's positions in the underlying circular and linear network respectively. In SW,

[17] Each agents was restricted to put weight at least equal to 20% to its own opinion. For the remaining 80%, a random number $\hat{t}_{i,j} \in (0,1)$ was generated for each link of the network (including the self link). The weight $t_{i,j} = 0.8\frac{\hat{t}_{i,j}}{\sum \hat{t}_{i,j}}$ for all $j \neq i$ with which i was connected and $t_{i,i} = 0.2 + 0.8\frac{\hat{t}_{i,i}}{\sum \hat{t}_{i,j}}$.

since the underlying network is circular, it does not matter who is indexed as 1st. In LSW, index 1 and 50 refer to the agents on the opposite edges of the underlying network. The vertical axis represents the average relative extremity of an agent's opinion in the first round (dashed line) and in the last round (solid line). A larger value implies larger average distance from the median, thus a more extreme opinion. This formulation entails less information than the opinion comparison matrix, yet for large networks this reduced form provides a concise picture. The two upper panels illustrate that when the network varies through a process that treats agents symmetrically in each round, like SF and SW, no patterns emerge regarding their relative positions, as expected. In contrast, the two lower panels show that additional structure in the way networks vary gives rise to distinct patters in agents relative positions. As shown in the lower left panel, in the FDSF simulations agents with a higher in-degree have tend to have more moderate opinions. In the lower right panel the LSW simulation results indicate that a more central (extreme) position in the underlying linear network would lead on average to more moderate (extreme) opinions.

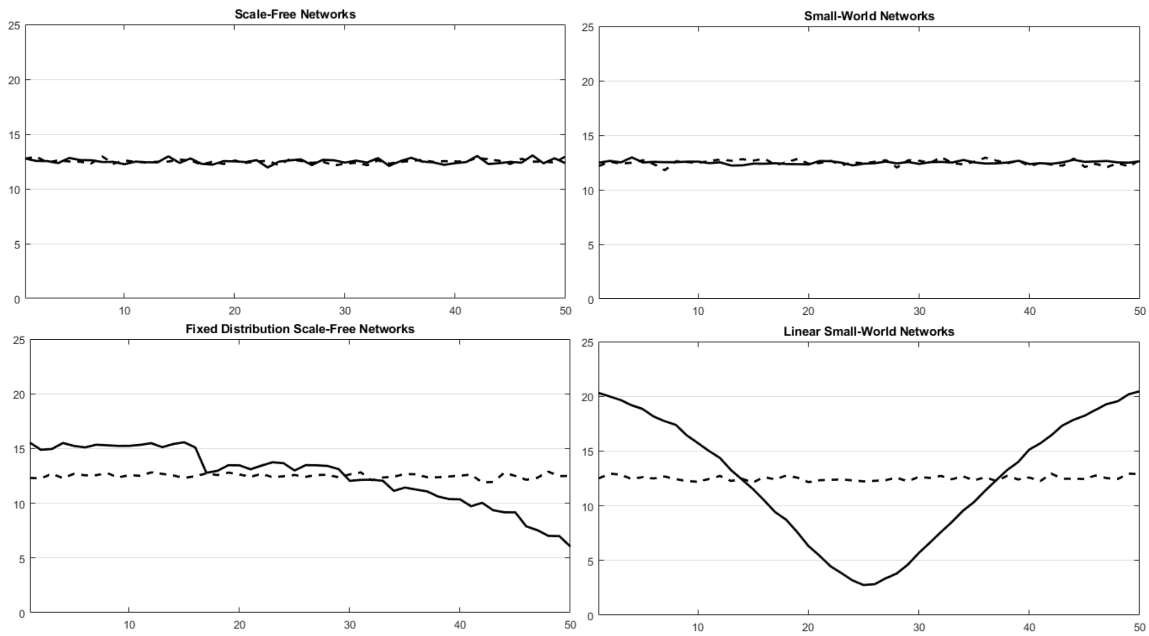


Figure 5: The average relative extremity of a node's opinion in the first round (dashed line) and in the last round (solid line) depending on the index of each node. In Fixed-Distribution Scale-Free networks nodes are indexed in increasing order of in-degrees (a higher index signifies a weakly higher degree). In Linear Small-World Networks nodes are indexed according to their position in the underlying linear network (1 is the leftmost node, 50 the rightmost). The average relative extremity is obtained by projecting opinions on the first principal component and calculating the absolute difference between a node's rank and the rank of the median, which is by construction 25.

Unfortunately, we cannot provide specific conditions under which an opinion updating process with varying networks gives rise to a unidimensional world. Still, our simulations demonstrate that a fixed communication network is not necessary for unidimensionality to arise. But some structure on communication, even if not necessarily a fixed network, is necessary for making predictions about relative positions of opinions. These simulations, jointly with our experimental results, pose what be believe to be interesting questions for future theoretical research regarding the necessary conditions that a sequence of listening matrices needs to satisfy in order to generalize the result of DeMarzo et al. (2003).

3 The Experiment

The theoretical predictions of DeMarzo et al. (2003) for fixed networks and our conjectures for the case of varying communication structures are tested in a lab experiment. We use three different experimental treatments. In all treatments, subjects repeatedly communicate their opinions about an unknown two-dimensional state. In two of the treatments, *Fixed 1* and *Fixed 2* the network remains fixed. In the third *Random* treatment, the network is varying, changing randomly in each round of communication, but keeping the in-degree of each node constant.

In what follows, we describe the way we induce and elicit opinions, how subjects communicate, and give more details about the network structures in the different treatments. We then formulate our experimental hypotheses.

3.1 Experimental Design and Procedures

The main task during the experiment is a non-competitive guessing game presented in the following form:

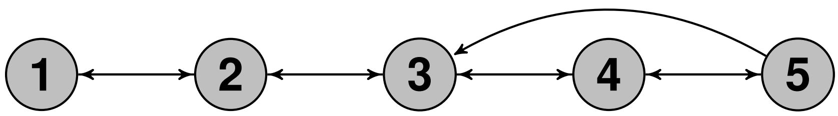
"In a tank there are 100000 balls. These balls are either RED or BLUE. The number of balls of each colour is random and any combination is equally likely. You are asked to guess the number of RED balls in the tank. This number could be anywhere between 0 and 100000."

The number of red balls represents an unknown state and the guess represents the subject's opinion about the state. Subjects play the guessing game for three phases, consisting of 10 rounds each.[18] In each phase subjects play two guessing games simultaneously over two different tanks: one with red and blue balls, and one with green and purple balls. The state and private signals for each of the two parallel games are drawn independently. In the first round, each individual observes a sample of 100 balls for each tank and makes two guesses. For each of the other rounds (2-10), subjects again make two guesses, observing both guesses for each of their neighbors (and no other signals). Thus, opinions are communicated over a two dimensional space. Subjects' experimental earnings depend on how close their guess is to the actual state.[19]

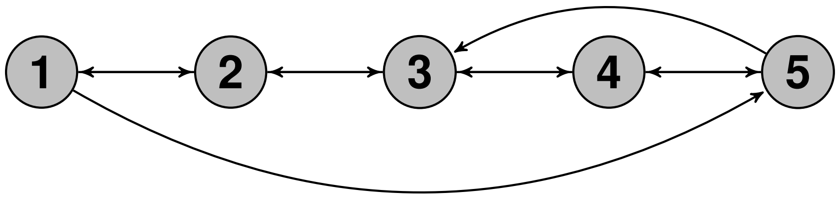
[18]The main experiment was preceded by two parts that aimed at familiarizing subjects with the information and communication environments. In part 1 subjects made guesses for a single tank without communication. In part 2 subjects could communicate but again made guesses for only one tank. Part 3 was the main part of the experiment described above. Instructions for all parts can be found in Appendix C.

[19]Guesses are incentivized using a quadratic scoring rule with a parameter that changes across rounds to

We use three **treatments:** ***Fixed 1*** **(F1),** ***Fixed 2*** **(F2)** and ***Random*** **(R)** as depicted in Figure 6. What varies across treatments is the network structure and its stability. In treatments *Fixed 1* and *Fixed 2* the network remains fixed but is different in each one (Figures 6a and 6b). *Fixed 1* serves as our baseline treatment. The network in *Fixed 2* is minimally different than the baseline: it is obtained by adding a single directed link to the baseline. In treatment *Random* the network structure changes randomly in each round of communication (Figure 6c). Each node observes the same number of neighbors as the corresponding node in the baseline. We explain the choice of the exact network structures at the end of this section, as it will be facilitated by the presentation of our research hypotheses.



(a) *Fixed 1*



(b) *Fixed 2*

(c) *Random*

Figure 6: Treatments. The graphs represent the 5-node network structure used for communication in each treatment. An arrow from one node to another means that the latter listens to the former. In treatment *Random* each node has the same number of neighbors as the corresponding node in *Fixed 1*, but these change randomly in each round.

In the two treatments where the network structure is fixed, the identity of each subject's neighbors (group members whose guesses are observed) remains fixed throughout the experiment and subjects are informed that this is so. Together, the sets of neighbors for each subject in a group form a directed network. Subjects are not informed about the structure of their group's network.[20] They are also told that observing another group member's guess

keep payoffs salient as opinions converge. Final earnings were determined by selecting randomly the payoffs in one of the three phases in part 1, one of the five rounds in part 2 and one of ten rounds for each of the three phases in the main part of the experiment in part 3. Subjects received an additional £3 participation fee.

[20]This as another difference in terms of design compared to related work on one dimension by Corazzini et al. (2012); Brandts et al. (2015); Battiston and Stanca (2015). We opted for an unknown network to guarantee

does not mean that that group member can observe their own guess. In the treatment with the varying network, subjects are informed that the number of neighbors they observe remains fixed, but a new set of neighbors is drawn randomly in each round. Neighbors are always drawn from within the same 5-member group.

Initial opinions in the experiment are induced by providing each subject at the beginning of each phase with a 100-ball sample from each tank. Each tank k contained a number θ_k of target balls (red for tank 1, green for tank 2). Each θ_k is drawn from a uniform distribution over $\{0, 1, 2, ..., 100000\}$. The samples were i.i.d. draws from a binomial distribution with parameters $n = 100$ and $p = \frac{\theta_k}{100000}$. For each group there was a set of $2 \times 3 = 6$ draws for θ_k (2 for each of the 3 phases) and $5 \times 2 \times 3 = 30$ samples (one for each of the 5 group-members in each phase). Across the experiment we used 3 such sets in approximately equal proportions in each treatment.

The experiment took place at the Lancaster Experimental Economics Laboratory hosted at the Department of Economics at the Lancaster University Management School (LUMS). A total of 180 subjects were recruited among LUMS students.[21] In total we had 12 groups for each treatment. Participants completed the tasks and interacted through computer terminals. The experiment was designed and run on z-Tree (Fischbacher, 2007). Average total payment was around £10.5 and the experiment lasted about 90 minutes.

3.2 Experimental Hypotheses

Following Theorem 1 and our conjectures supported by simulations, we expect *Unidimensionality* to arise in all treatments, as long as subjects update their guesses for both dimensions by taking some form of average of their own and others' guesses of the previous round. Whether the network is fixed or varying should not matter. We therefore consider the following hypothesis:

Hypothesis 1. *The variance of guesses explained by the first principal component converges to 1 in all treatments.*

Theorem 1 suggests that if in each fixed network groups use the same listening matrix and this differs between the two networks, then we should expect differences in subjects' relative positions. Given that our simulations suggest that this will be true also for the varying network, we expect relative positions to differ across all treatments.

comparability between the designs of fixed and random networks where in the latter keeping track of the continuously changing network would be practically infeasible. Notice that i) the theory of DeMarzo et al. (2003) does not require knowledge of the network structure, and ii) Grimm and Mengel (2020) precisely show that when the network is unknown, naive models of updating have more bite than when the network is common knowledge.

[21]There were 15 sessions in total. Most sessions had 15 participants, except for 1 with 10 subjects and 3 with 5 subjects due to low turnout. Initially 190 students were recruited. During one of the sessions a technical issue affected the play of two groups (10 subjects) altering the intended treatment. We excluded this data from further analysis.

Hypothesis 2 (a). *Relative positions as projected on to the long-run opinions' first principal component should differ across all treatments.*

To obtain more crisp predictions about subjects' relative positions after some rounds of communication we focus on the theoretical predictions assuming that subjects assign an equal weight to their own and every other neighbor's opinion for each round of communication. While strong, this assumption seems the most natural as a theoretical benchmark, since subjects in the experiment are not aware of the network structure and therefore all neighbors are identical to them. Updating according to (3) then requires $\lambda(t) = 1$ for all t and listening matrices $\mathbf{T}_{F1}$ and $\mathbf{T}_{F2}$ defined as follows:

$$\mathbf{T}_{F1} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \\ 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix} \qquad \mathbf{T}_{F2} = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 \\ 0 & 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 1/3 & 1/3 & 1/3 \\ 1/3 & 0 & 0 & 1/3 & 1/3 \end{pmatrix}$$

The difference between these two listening matrices is in the last row representing the weight that agent 5 assigns to their neighbors. While in *Fixed 1* agent 5 is only observing the guesses of agent 4, in *Fixed 2* agent 5 is also observing the guesses of agent 1.

Following DeMarzo et al. (2003) and due to the position determinacy property (Theorem 1), relative positions in the fixed treatments are determined by the ranking of agents' corresponding element in the second column eigenvector of $\mathbf{T}$. For $\mathbf{T}_{F1}$ this is equal to $V_2^c(T_1) = (-2.5, -1.667, 0, 0.667, 1)^T$, which means that the relative positions for the five agents from extreme left to extreme right is $(1, 2, 3, 4, 5)$ –or $(5, 4, 3, 2, 1)$. For $\mathbf{T}_{F2}$ the second column eigenvector is equal to $V_2^c(T_2) = (-4.5, -1.068, 3.585, 5.356, 1)^T$, which means that the relative positions of the five agents from extreme left to extreme right is $(1, 2, 5, 3, 4)$ –or $(4, 3, 5, 2, 1)$.[22]

In the *Random* treatment the updating matrix can not remain the same and Theorem 1 does not apply. We therefore base our hypothesis on our conjecture, which is also supported by simulations. The simulation using the Fixed Distribution Small-World network formation process is particularly relevant here. Similar to that, in the *Random* treatment each individual observes a constant number of neighbors. We expect individuals with the highest in-degree to be more central, while the lowest in-degree to be more often in the extremes. Using simulations and assuming equal weights we can calculate the expected opinion comparison matrix $E\left[\hat{\mathbf{C}}^{\kappa,\lambda}_{random}\right]$. Together with the predictions via strong position determinacy (for *Fixed 1* and *Fixed 2*), we summarize the above as follows:

Hypothesis 2 (b). *Assuming that all subjects in a given treatment update their opinions by assigning equal weight to their own and their neigbors' opinions, relative positions as projected on to the long-run opinions' first principal component will converge to the following opinion comparison*

[22]Notice that these are also the relative positions of individuals in the examples of unidimensional worlds shown in Figure 3: in panel B for $\mathbf{T}_{F1}$ and in panel C for $\mathbf{T}_{F2}$.

matrices:

$$\hat{\mathbf{C}}^{1,3}_{fixed\ 1} = \begin{pmatrix} . & . & . & . & . \\ 0 & . & . & . & . \\ . & 0 & . & . & . \\ 0 & 0 & 0 & . & . \\ 0 & 0 & 0 & 0 & . \end{pmatrix} \qquad \hat{\mathbf{C}}^{1,3}_{fixed\ 2} = \begin{pmatrix} . & . & . & . & . \\ 0 & . & . & . & . \\ . & 0 & . & . & . \\ 0 & 0 & 0 & . & . \\ 0 & 0 & 1 & 1 & . \end{pmatrix}$$

$$E\left[\hat{\mathbf{C}}^{1,3}_{random}\right] = \begin{pmatrix} . & . & . & . & . \\ 0.18 & . & . & . & . \\ . & 0.46 & . & . & . \\ 0.18 & 0.46 & 0.55 & . & . \\ 0.21 & 0.44 & 0.45 & 0.43 & . \end{pmatrix}$$

Notice that we present the opinion comparison matrices fixing $\kappa = 1$ and $\lambda = 3$. The predicted matrices for *Fixed 1* and *Fixed 2* are not affected by this choice, but we would get slightly different numbers for the prediction for *Random*. Such choice guarantees that subjects at these nodes in the two fixed networks remain at the same distance in both treatments and are the furthest apart from all such pairs of nodes hence making the exercise less sensitive to noise. Also notice that for illustrative purposes, we present only the nine elements of each matrix that are sufficient to completely describe all subjects' relative opinions.[23]

Having presented our hypotheses, we can now explain the choice of the particular network structures for our experiment. For the baseline structure (*Fixed 1*) we had two desiderata: i) it should be simple so that theoretical predictions regarding subjects' relative positions in the opinion space can be directly traced back to their position in the network; ii) it should be possible to minimally alter the baseline and obtain a substantially different prediction for relative opinions, in order to test our hypothesis of the direct link between the ordering and the network structure. A simple undirected linear network would satisfy the first condition but not the second: adding a single directed link to the line, like in *Fixed 1*, gives the same predicted ordering as the starting line. On the contrary, by adding to *Fixed 1* a single directed link from node 1 to node 5 we obtain *Fixed 2*, where predicted relative positions are now different. This satisfies our second condition, providing the desired test for our ordering hypothesis.

4 Updating in the lab

For illustration of our data, Figure 7 shows the evolution of opinions in a phase of ten rounds for one of the groups in the experiment. There is of course heterogeneity in our data, but this figure facilitates the understanding of the different measures we use to summarize the data.

[23] That is, we exclude the diagonal and (3,1) elements that are zero by definition and all entries above the diagonal given that the matrix is anti-symmetric.

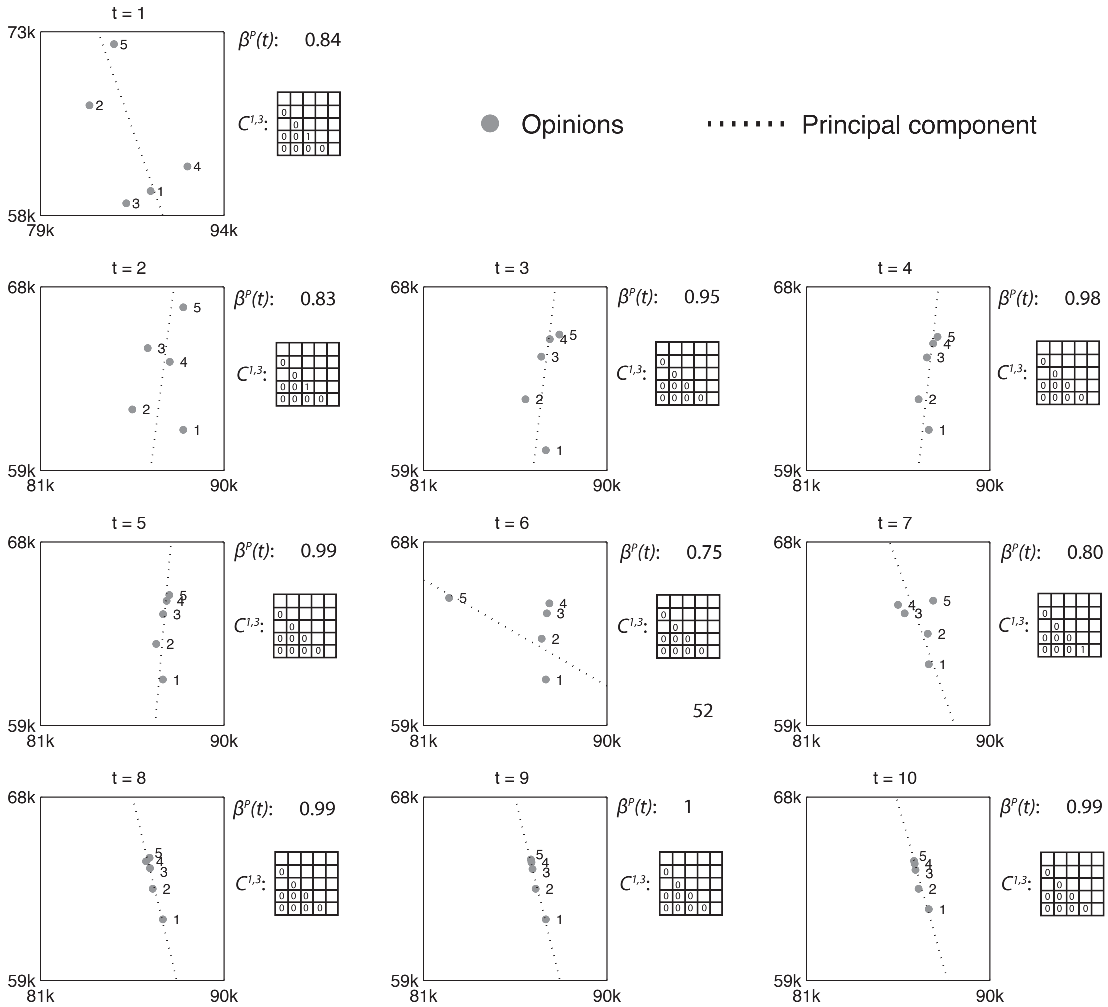


Figure 7: Guesses across rounds in a phase for a particular group in treatment *Fixed 1* in the experiment (*session 6, group 2, phase 2)*. Each graph represents guesses in each round t, from 1 to 10. Each point represents a subject's guesses (in thousands) for each tank. Tank 1 in the horizontal axis and tank 2 in the vertical axis. Labels 1 to 5 refer to subjects positions in the network. The dashed line traces the first principal component. $\beta^P(t)$ is the variance explained by the first principal component in t. The opinion comparison matrix is $C^{1,3}$.

First of all, guesses get closer over time and most of this convergence happens in the first rounds. More importantly, we also see that guesses quickly become unidimensional, captured by the high percentage of variance explained by the first principal component, $\beta^P(t)$. The relative positions of subjects' projection on this line also converges quickly and follows the order of their labels: 1 next to 2, next to 3 etc. This is captured by the opinion comparison matrix $C^{1,3}$. The process of convergence, as captured by all these different measures is interrupted in round 6, where subject 5 makes a guess for tank 1 that is far from their own and others' previous guesses. Such "jumps" or perturbations occur sometimes in the data. Interestingly, we observe that the process of convergence to unidimensionality picks up again immediately, only now on a line 'tilted' by the jump.

4.1 Unidimensionality

We now turn our attention to the first of our main research questions: can communication lead to unidimensional opinions? Recall that based on Theorem 1 and our simulations, this should be true for an arbitrary sequence of listening matrices and hence, as stated in Hypothesis 1, we should observe evidence of this in all three experimental treatments.

The left panel of Figure 8 depicts the mean percentage of variance explained by the first principal component $\beta^P(t)$ in our data, for each round t in each treatment. Note that with two dimensions by definition $\beta^P(t) \in [50, 100]$. The middle panel presents the benchmark theoretical prediction of our model if agents were to assign equal weights to their own and their neighbors' opinions (see $\mathbf{T}_{F1}$ and $\mathbf{T}_{F2}$). It is clear that in the theoretical benchmark $\beta^P(t)$ approaches 100% as theory dictates. In our data, despite not observing so high values for the average there is an increase in $\beta^P(t)$ across all treatments.[24]

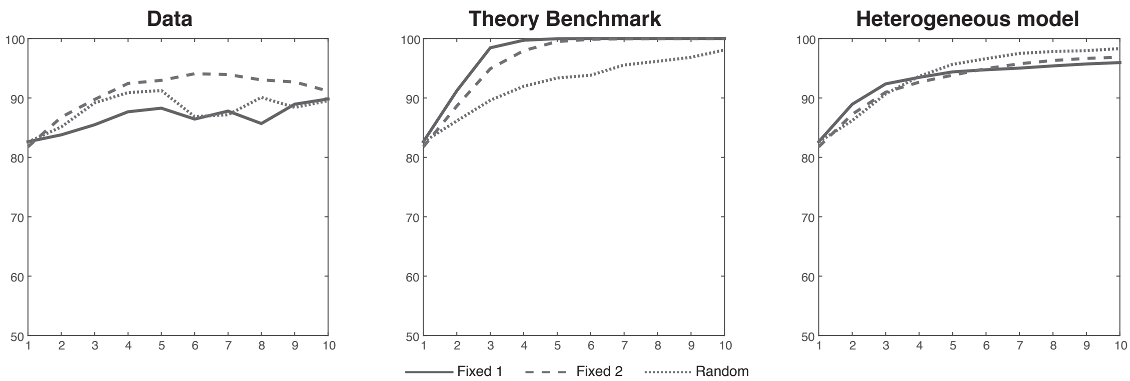


Figure 8: Mean percentage of explained variance across round and treatment in a) our data (left), b) theoretical predictions assuming all subjects behave the same (middle), c) theoretical predictions assuming an heterogeneous behavior (right).

Result 1. Hypothesis 1, stating that in all treatments the variance of guesses explained by the first principal component converges to 1, cannot be rejected.

Support: Figure 8 provides some graphical support for this result based on aggregate data. More conclusively, based on the non-parametric seasonal Mann-Kendall test for trend, we can reject the hypothesis that there is no positive trend in the series for β^P in *Fixed 1* ($p < 0.001$), in *Fixed 2* ($p < 0.001$) and in *Random* ($p = 0.002$).[25] Furthermore, the mean (median) value of $\beta^P(t)$ in rounds 6 to 10 is 87.4 (90.4) in *Fixed 1*, 93 (96.9) for *Fixed 2* and 88.4 (92.3) for *Random*.

There are a few things to note with respect to this result. First of all, while convergence to unidimensionality holds for all treatments, it is noticeably stronger in *Fixed 2*. As it turns out, in this treatment there is less noise due to perturbations, which could explain this difference.[26] Second, the graphs indicate that the increase in β^P happens mostly in the

[24] We defer discussion of the *Heterogeneous model* on the right panel after having presented our main results.

[25] Note that the seasonal Mann Kendall test uses information from individual group observations, not just the aggregate data shown in Figure 8.

[26] This is discussed in more details in Appendix A.

first five rounds. This is confirmed by running the seasonal Mann-Kendall test for trend but restricting the data to the respective rounds. For rounds 1 to 5 we can reject the null that there is no positive trend in all three treatments ($p < 0.001$ for all three tests). For rounds 6 to 10 we cannot reject the null of no trend in any of the treatments ($p = 0.484$ in *Fixed 1*, $p = 0.134$ in *Fixed 2*, $p = 0.388$ in *Random*). This "early action" in our data is in line with most updating taking place in the first rounds. As we show in Appendix A, convergence of opinions as measured by the normalized coefficient of variation also mainly occurs in the first 5 rounds. Third, the mean β^P in round 1 seems relatively high, which can lead to the concern that this may explain the high levels achieved in subsequent rounds. Recall that in this round subjects have not yet observed any of their neighbors guesses. The high β^P can therefore only be attributed to the random draw of private signals used in the experiment. Following standard experimental procedures, the draw was kept random, as explained in the instructions, and we did not make any selection of specific signal sets. More to the point, there is no significant correlation between a group's $\beta^P(1)$ and its average β^P for the last 6 rounds. This means that the increasing trend we observe is not a result of the (on average) high β^P induced by the initial draw of signals.

4.2 Relative Positions

To compare subjects' relative positions we rely on the *opinion comparison matrices*, $C^{1,3}_{i,j}$. Recall that by definition, any matrix $C^{\kappa,\lambda}_{i,j}$ is anti-symmetric, the elements of the diagonal are 0, and the elements (κ, λ) and (λ, κ) are 1 and 0 respectively. It therefore suffices to look at the remaining elements below the diagonal to have a complete picture of the relative positions in the group. For the groups in our experiments these are the following 9 elements: (2,1), (3,2), (4,1), (4,2), (4,3), (5,1), (5,2), (5,3), (5,4).

The left panel of Figure 9 shows the average value for each of those elements across rounds for all observations in our data for each treatment. Taking into account subjects' initial guesses, the middle panel presents the theoretical predictions of Hypothesis 2(b), with all subjects updating their opinions in the same manner and putting equal weight to their own and their neighbors' opinions. Observing Figure 9, we see that all elements in *Fixed 1* converge to average values below 0.5. In *Fixed 2* the same is true for all elements except (5,3) and (5,4). Recall from Hypothesis 2(b) -and the panel in the middle- precisely these two elements should contain a treatment effect across the two fixed network treatments. This is due to the expected change in the predicted position of individual 5 being moving from the extreme position on the line in *Fixed 1* to a moderate position in *Fixed 2*. The exact prediction according to Hypothesis 2(b) would be for these two elements to converge to 1, which we do not observe in our data. Still, their average values do remain above 0.5. In *Random* we do not observe any tendency for elements to converge to extreme values. In fact, on average elements take values very close to the prediction obtained in Hypothesis 2(b).

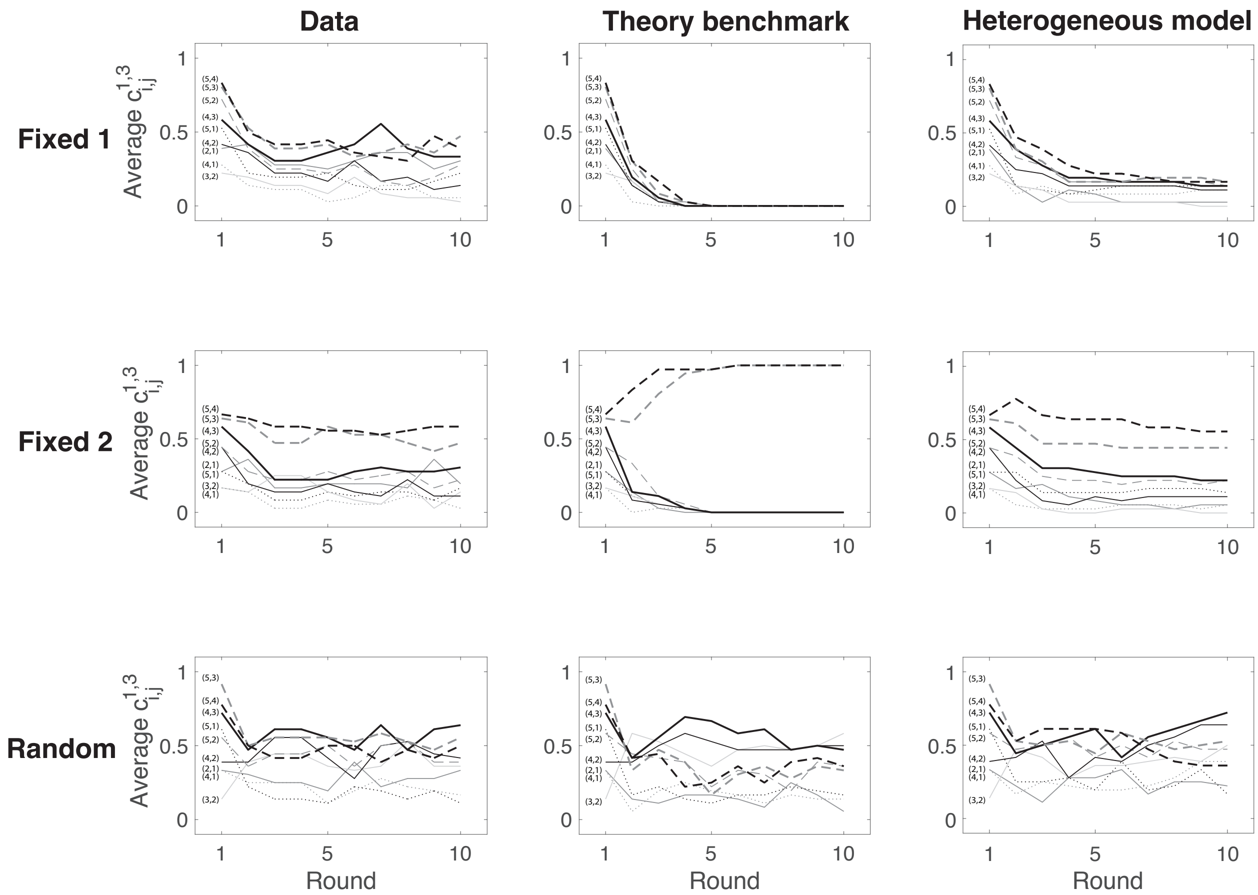


Figure 9: Each line corresponds to an element below the diagonal of the opinion comparison matrix, specified by the pairs in parenthesis on the left of each graph. Values indicate the average value for that element across all groups in each treatment in a given round. The graphs on the left show the values in our experimental data. The graphs on the middle simulate how the initial guesses in our data would evolve if all subjects behave the same and put equal weights to all their neighbors and their own previous guess as in Hypothesis 2(b). The graphs on the right simulate how the initial guesses in our data would evolve permitting an heterogeneous behavior among subjects.

Result 2. Hypothesis 2(a) can not be rejected, Hypothesis 2(b) is partially supported.

Support: Concerning Hypothesis 2(a) we compare the data appearing in the first row and column of Figure 9, with the data right below it. In particular, we compare the average values of elements in $C^{1,3}_{Fixed\,1}$ and $C^{1,3}_{Fixed\,2}$, averaged across rounds 6 to 10. A Fisher's exact test returns significant or slightly significant differences between elements (2,1) ($p = 0.024$), (4,3) ($p = 0.027$), (5,3) ($p = 0.072$) and (5,4) ($p < 0.001$). We therefore conclude that relative positions in these two treatments are different. Making the same comparison between elements in $C^{1,3}_{Random}$ and the in other two matrices we obtain significant or slightly significant differences between most elements. Element (5,1) is the only one that does not display significant differences across all treatments.

Concerning Hypothesis 2(b) we compare the data in the three graphs in the first column of Figure 9 with those in the second column. In particular, we compare the average values in $C^{1,3}_{Fixed\,1}$, $C^{1,3}_{Fixed\,2}$ and $C^{1,3}_{Random}$, averaged across rounds 6 to 10, with the corresponding theoretical predictions from Hypothesis 2(b). For *Random* we only find significant differences for element (2,1) ($p = 0.014$): the value in the data is 0.30 when theory predicts 0.18. For both *Fixed 1* and *Fixed 2* all elements of the empirical comparison matrices differ significantly from the theoretical predictions. Nevertheless, as mentioned in the discussion above, some of the qualitative features of the theoretical predictions do manifest in the data.

Figure 10 provides an alternative view of the results regarding relative positions in the *Random* treatment that allows for a better comparison with our simulation results presented in Figure 5. Instead of focusing on the comparison matrix, we focus on individuals' average extremity. Recall that the simulation using the Fixed Distribution Small-World network formation supported that individuals with high (low) in-degree are expected to form moderate (extreme) positions. This is precisely what we observe in Figure 10 where on the left we present the empirical patterns and in the middle the theoretical benchmark predictions. Here individuals are indexed as in the presentation of our treatments in Figure 6 where individuals 1 and 5 observe the opinions of one neighbour, individuals 2 and 4 those of two neighbours, and individual 3 that of of 3 neighbours. Despite starting the process with the average extremity values induced by the random signals (dashed line) being uncorrelated to the number of neighbors observed, we find that after updating individuals' extremeness is decreasing in their in-degree.

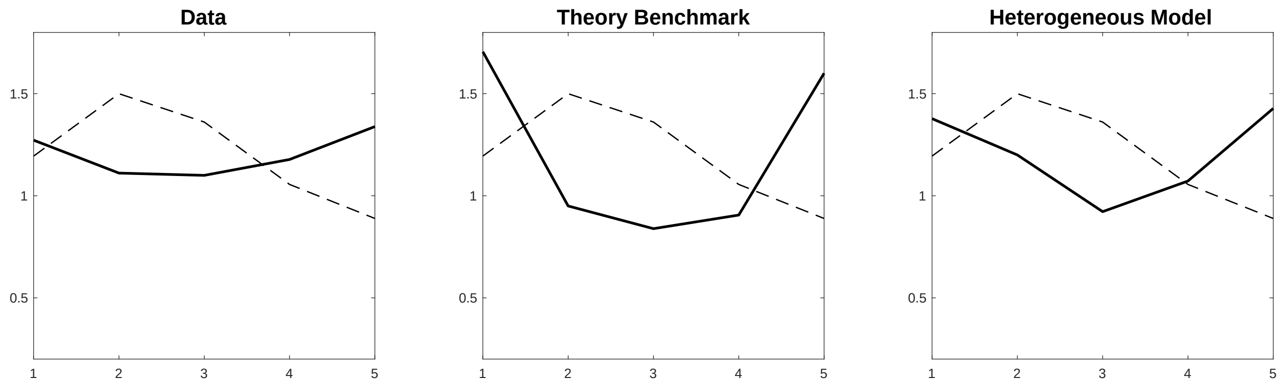


Figure 10: The average extremity of a node's opinion in the first round (dashed line) and in rounds 6 to 10 (solid line) depending on the index of each node for treatment *Random*. The average extremity is obtained by projecting opinions on the first principal component and calculationg the absolute difference between a node's rank, taking values from 1 to 5 and the median rank, which is 3.

4.3 Heterogeneity

As we show in the previous section, when comparing our experimental results to the theoretical benchmark we find evidence in support of the predictions about unidimensionality put forth by De Marzo et al. (2003) and our simulations. As in any empirical exercise, we observe some divergence from the theoretical bencharks. In this section we ask whether such divergence can be accommodated within the framework of the averaging model, or if this framework is inadequate to do so. We show how a simple extension of the homogeneous averaging model introduced earlier, which allows for heterogeneity among individuals' updating behavior, can account for most of the divergence we find and, when simulated, traces our experimental observations well.

The theoretical benchmark we used for the design of the experiment and the formulation of our hypotheses is given in equation (3). In that "homogeneous" model the updating behavior is captured by $\lambda(t) \in (0, 1]$. Notice that to obtain our theoretical benchmark we used $\lambda(t) = 1$ for all t. Using different updating types (λ functions) could offset some of the divergence of the data from the theory to a limited extend. Here we go a step further allowing different individuals $i \in D$ to have different updating behaviors, captured by a respective λ function. Formally equation (3) is replaced by:

$$T_{i,j}(t) = (1 - \lambda_i(t))\mathbf{1}_{i=j}(j) + \lambda_i(t)\frac{1}{|D_i(t) \cup i|} \tag{4}$$

All else remains as before. A value of $\lambda_i(t) = 1$ means that the individual places equal weight to all neighbors' opinions including their own. A value of $\lambda_i(t) = 0$ means that an individual keeps the same opinion as in the previous round and does not perform any updating.

To maintain some of the model's parsimony while allowing for heterogeneity we turn to the data. We use cluster analysis (k-means clustering) to identify a limited number of patterns of updating behavior.[27] Figure 11 shows the updating types identified. The type represented by a horizontal line at 1 is essentially one that takes an exact average of their neighbors' and their own opinion in every round. This coincides with the behavior assumed for the theoretical benchmark. The type represented by the line starting near 0.3 and dropping close to 0 within a few rounds, represents an individual that only updates their initial opinion slightly in the first couple of rounds and essentially stops updating after that. Note that this behavior is not allowed in the DeMarzo et al. (2003) model on which we built on. In particular it violates the condition $\sum_{t=1}^{\infty} \lambda(t) = +\infty$. The updating patterns of the other two types lie in between these two extremes. It is worth noting that the two described types of "perfect averaging" and "stubborn" subjects come up consistently in the cluster analysis exercise, independently of the choice of parameters used.

Having obtained these four type of individuals we next simulate the model using subjects' initial guesses as $\mathbf{s}(1)$. The results of these simulations in terms of the percentage of variance explained by the principal component and the relative positions are shown in the respective panels of Figures 8 and 9.

The right panel in Figure 8 shows the mean percentage of explained variance for each treatment obtained in the simulations of the heterogeneous model. In the graph we see that in contrast to the benchmark homogeneous model, the measure does not converge to 100% for any treatment. In fact, after the first five rounds it reaches levels that are only slightly

[27] The exact way in which this is done is described in Appendix B. The codes used are available from the authors upon request.

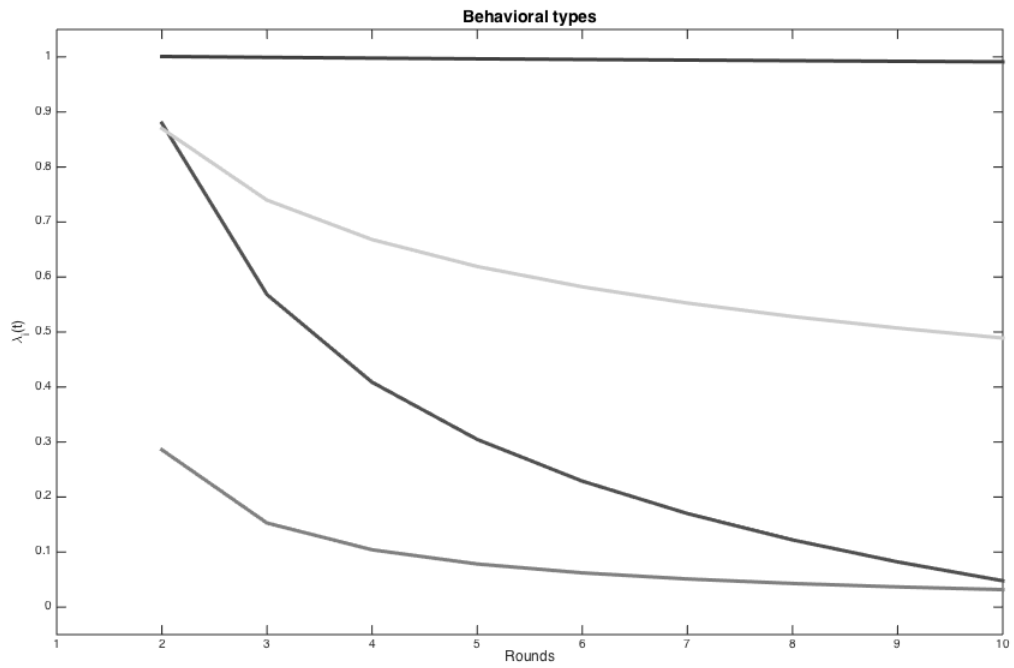


Figure 11: Updating types identified through cluster analysis and used in the Heterogeneous model. The graph illustrates the weight $\lambda_i(t)$ assigned by individuals of each of the four identified types across the ten rounds.

higher than the ones in the data for treatments *Fixed 2* and *Random* (Recall, that we observe most updating happening in the first five rounds. See Appendix A).

We obtain similar findings regarding individuals' relative positions. The right panels of Figure 9 show how the average relative positions evolve in each treatment in the simulations of the heterogeneous model. Perhaps the most striking result here concerns treatment *Fixed 2*. Under the homogeneous model, elements (5,4) and (5,3) should converge to one and all others to zero. This distinction is only qualitatively reflected in the experimental data, where these two elements do obtain higher values, but still remain far from one. On the other hand, in the heterogeneous model these two elements reach very similar levels. All other elements obtain levels closer to zero, but still do not reach zero, which is exactly what happens in the data. In the *Fixed 1* and *Random* treatments we can also observe that the heterogeneous model goes a long way reconciling the predictions of the averaging model with our experimental results.

The potential downside of our approach is that of "over-fitting". As this exercise is mainly exploratory we do not intend to draw inferences about the empirical relevance of the identified updating types outside the domain of this experiment. The main objective is to obtain some feeling of the degree to which simply introducing heterogeneity into the DeMarzo et al. (2003) framework can explain the divergence of the data from the theoretical benchmark. Having said that, we do integrate some regularization techniques into the cluster analysis to mitigate over-fitting (see Appendix B). Still, our main result, namely that the heterogeneous model explains much of the divergence between the data and the theoretical benchmark, is robust to any reasonable choice of parameters.[28]

5 Conclusions

Since DeGroot (1974), the simple averaging model has been used extensively to model opinion formation. Recent experiments have given empirical support to its behavioral assumptions. This solid ground, makes the model's predictions even more relevant. In particular, the finding by DeMarzo et al. (2003) about the emergence of a unidimensional world in fixed communication networks. In this paper, our similations and experiments cast a conjecture on the generalizability of their result, showing the robustness of unidimensionality even when communication networks are not fixed. At the same time we highlight the role of the network in determing the characteristics of the unidimensional world that emerges. Our experiment empirically validates the theoretical findings, and our analysis shows that a simple extension that allows for heterogeneity permits the simple averaging model to closely trace the experimental data.

Naturally, our result is obtained in a stylized environment governed by several assumptions. One we consider most restrictive, is that individuals use the same listening matrix to update opinions across all issues. This implies that at any point in time, individuals listen the same set of individuals for all issues and assign the same weight to a given individual's opinion for all issues. In our experiment, this may not be an issue due to its stylized environment. In natural settings however, it is reasonable to think that individuals may place different weights on a friend's opinions about political issues and sports. In fact, one may only discuss specific issues in certain social circles. It remains hence an open question how different the listening matrices for each dimension can be to still observe the emergence of unidimensional worlds. The above, as well as the robustness of the unidimensionality result in settings where agents do not passively hear their neighbors' opinions, but instead actively engage in shaping their communication networks as in recent research (e.g., Sethi and Yildiz, 2016; Melguizo, 2019) may provide fruitful avenues for future research.

[28]The interested reader may use our code to explore how some of the parameter choices involved in the cluster analysis exercise may result in identifying different updating types.

References

Ansolabehere, S., J. Rodden, and J. M. Snyder (2008): "The strength of issues: Using multiple measures to gauge preference stability, ideological constraint, and issue voting," *American Political Science Review*, 102, 215–232.

Barabási, A.-L. and R. Albert (1999): "Emergence of scaling in random networks," *Science*, 286, 509–512.

Battiston, P. and L. Stanca (2015): "Boundedly rational opinion dynamics in directed social networks: Theory and experimental evidence," *Journal of Economic Behavior and Organization*, 119, 400–421.

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

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

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

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

DellaPosta, D., Y. Shi, and M. Macy (2015): "Why do liberals drink lattes?" *American Journal of Sociology*, 120, 1473–1511.

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

Enke, B. and F. Zimmermann (2019): "Correlation neglect in belief formation," *The Review of Economic Studies*, 86, 313–332.

Fischbacher, U. (2007): "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 10, 171–178.

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

——— (2012): "How homophily affects the speed of learning and best response dynamics," *Quarterly Journal of Economics*, 127, 1287–1338.

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

Jackson, M. O. (2008): *Social and economic networks*, Princeton University Press, Princeton.

Kosinski, M., D. Stillwell, and T. Graepel (2013): "Private traits and attributes are predictable from digital records of human behavior," *Proceedings of the National Academy of Sciences*, 110, 5802–5805.

Lee, D. S., E. Moretti, and M. J. Butler (2004): "Do voters affect or elect policies? Evidence from the U.S. house," *Quarterly Journal of Economics*, 119, 807–859.

Levy, G. and R. Razin (2015): "Correlation neglect, voting behavior, and information aggregation," *American Economic Review*, 105, 1634–1645.

——— (2016): "Does polarization of opinions lead to polarization of platforms? The case of correlation neglect," *Quarterly Journal of Political Science*, 10, 321–355.

——— (2018): "Information diffusion in networks with the Bayesian peer influence heuristic," *Games and Economic Behavior*, 109, 262–270.

——— (2019): "Echo chambers and their effects on economic and political outcomes," *Annual Review of Economics*, 11, 303–328.

McMurray, J. (2014): "Why the political world is flat: An endogenous left-right spectrum in multidimensional political conflict," in *APSA 2014 Annual Meeting Paper*.

Melguizo, I. (2019): "Homophily and the Persistence of Disagreement," *The Economic Journal*, 129, 1400–1424.

Molavi, P., A. Tahbaz-Salehi, and A. Jadbabaie (2018): "A Theory of Non-Bayesian Social Learning," *Econometrica*, 86, 445–490.

Nedic, A. and J. Liu (2014): "A Lyapunov approach to discrete-time linear consensus," in *Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference on*, IEEE, 842–846.

Nedic, A. and A. Ozdaglar (2009): "Distributed subgradient methods for multi-agent optimization," *IEEE Transactions on Automatic Control*, 54, 48–61.

Ortoleva, P. and E. Snowberg (2015): "Overconfidence in political behavior," *American Economic Review*, 105, 504–535.

Page, L. (2020): "Assessing the unidimensionality of political opinions. An indirect test of the persuasion bias," *Journal of Economic Behavior & Organization*, 174, 469–477.

Poole, K. T. and H. Rosenthal (1997): *Congress: A political economic history of roll–call voting*, Oxford University Press, Oxford, UK.

Sethi, R. and M. Yildiz (2016): "Communication with unknown perspectives," *Econometrica*, 84, 2029–2069.

Siegel, S. (1961): "Decision making and learning under varying conditions of reinforcement," *Annals of the New York Academy of Sciences*, 89, 766–783.

Spector, D. (2000): "Rational debate and one–dimensional conflict," *Quarterly Journal of Economics*, 115, 181–200.

Tahbaz-Salehi, A. and A. Jadbabaie (2008): "A necessary and sufficient condition for consensus over random networks," *IEEE Transactions on Automatic Control*, 53, 791–795.

Watts, D. J. and S. H. Strogatz (1998): "Collective dynamics of 'small-world'networks," *Nature*, 393, 440–442.

Wilson, C. and J. Haidt (2014): "It's true: Liberals like cats more than conservatives do," *TIME, published on Feb. 18th 2014*.

Appendix

A Updating, convergence and perturbations

Notice that two main assumptions in our theoretical model are that a) agents update their opinions (in a way that resembles averaging over one's neighbors' opinions), and b) they do this in the same way for both dimensions. If the former holds then opinions need to come closer over time, something that would be reflected on a diminishing normalized coefficient of variation (NCV).[29] If the latter holds, then the decrease of the NCV should be the same across the two dimensions.

The three upper panels of Figure 12 show the average NCV per round of guesses for each tank in each treatment. What immediately stands out is the jump in the value for tank 2 in round 3 of treatment *Fixed 1* that jumps up to 150%. This is mostly driven by a particular case where the group's NCV for tank 2 jumped up to 2653%. Jumps like that, although smaller in magnitude are common in the data. See the example in Figure 7 for such an instance in $t = 6$. Some of those, especially the biggest in magnitude, can be attributed to 'mistakes', such as mis-typing one's guess. Others may be deliberate, although there does not seem to be some systematic pattern of behavior to explain them.[30] We do observe that such perturbations are less common in *Fixed 2*. This can be seen in Figure 12 by noticing that the mean and median for the NVC in *Fixed 2* are closer than in the other two, for both dimensions, which is evidence of a distribution with fewer extreme values.

Irrespective of what causes them, these perturbations can be useful in our study. Individuals' opinions in real life may also be subject to shocks. Even if their causes are different from what makes subjects in the lab "jump", having this feature in the lab can be informative about the robustness of the unidimensionality properties to similar "noise".

The lower panels of Figure 12 show again the median NCV in each round, for each tank, in all three treatments. The pattern of convergence can be seen much better here. We do not observe any systematic differences in the convergence pattern between the two dimensions in each treatment.[31] Across treatments we observe that convergence appears to last longer in *Random*, where it also reaches higher levels (lower NCV). In all treatments, we see that NCV decreases mostly in the first five rounds and remains rather flat in the last five rounds.[32]

[29]The coefficient of variation of guesses for one tank is $CV(t) = \frac{st.dev.(guesses\ at\ t)}{mean(guesses\ at\ t)}$. We then report the normalized coefficient $NCV(t) = \frac{CV(t)}{CV(1)}$.

[30]One reason for such perturbations could be that a subject tries to hedge by making guesses across a reasonable range of values, even if that is not optimal. Another reason could be boredom, which can affect subjects' behavior in repetitive experiments as this and may lead to arbitrary choices. The observation that boredom may lead to random responses in guessing games was first made by Siegel (1961).

[31]A Wilcoxon signed-rank test comparing pairs of groups' NCV for each dimension in each round of each treatment does not reject the null that these are different in any but two instances: in round 7 of treatment *Fixed 2* ($p = 0.053$) and in round 5 of treatment *Random* ($p = 0.04$). Given the high number of comparisons (3 treatments $\times$ 9 rounds = 27 comparisons) it is expected to obtain some false positives. Applying any of the standard corrections for multiple testing will render both these cases non-significant even at the 10% level.

[32]Based on the non-parametric seasonal Mann-Kendall test for trend, we can reject the hypothesis that there is no trend in the coefficient of variation series for guesses for Tank 1 ($p < 0.001$) and Tank 2 ($p < 0.001$) in *Fixed 1*, for Tank 1 ($p < 0.001$) and Tank 2 ($p < 0.001$) in *Fixed 2*, and for Tank 1 ($p < 0.001$) and Tank 2 ($p < 0.001$) in *Random*.

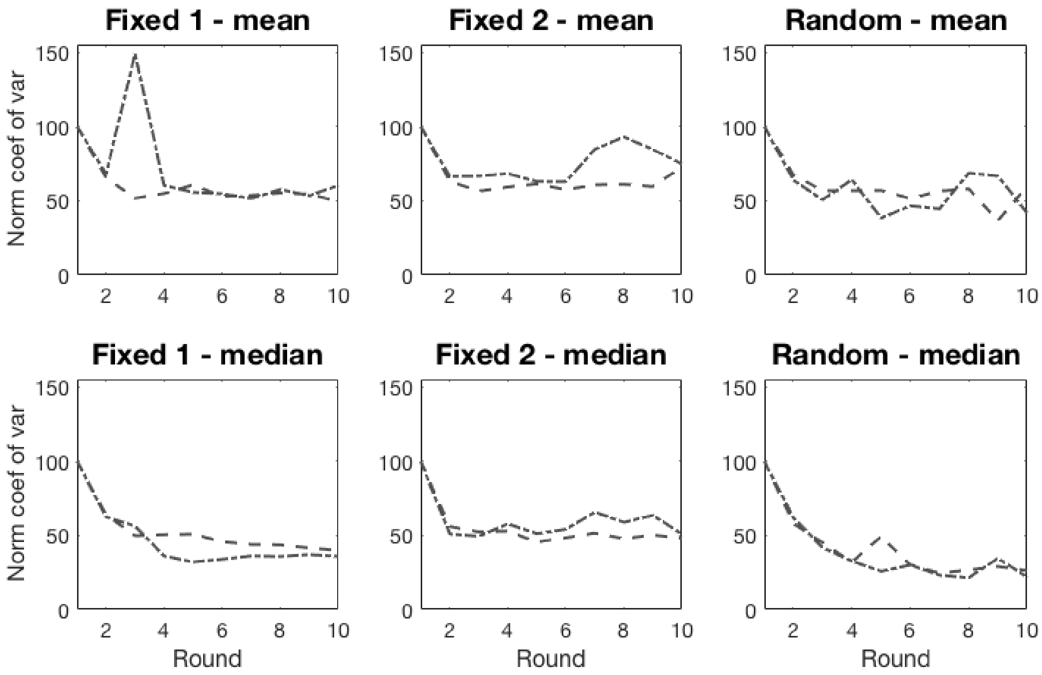


Figure 12: The graphs show the mean and median NCV values per round in each treatment. *Dashed lines* correspond to the NCV of guesses for *tank 1* and *dot-dashed lines* correspond to the same for *tank 2*.

B Clustering analysis

Combining equations (1) and (4) gives us an individual's opinion on issue k at time $t+1$ as:

$$s_i^k(t+1) = (1-\lambda_i(t))s_i^k(t) + \lambda_i(t)\frac{1}{|D_i(t)|}\sum_{j\in D_i}[s_j^k(t)]$$

.

Each subject played 3 phases x 10 guesses for tanks 1 and 2. For each individual and each of these guesses we solve this equation for $\lambda_i(t)$ and obtain a $\hat{\lambda}_i^k(t_p)$, where $p \in \{1,2,3\}$ refers to the phase. As subjects guesses are not restricted, it happens that some of these values lie outside of the theoretical bounds: $0 \leq \lambda_i(t) \leq 1$. While this does happen, most values obtained are within or very close to the theoretical bounds. To avoid extreme values biasing the excercise, we set a censoring parameter $c = .3$ and apply the following set of rules: if $\hat{\lambda}_i^k(t_p) < -c$ or $\hat{\lambda}_i^k(t_p) > 1+c$ for one of the tanks but not the other, we only use the moderate value. If both values are extreme we consider them as missing. If none of the values are extreme, we take the average of the two: $\lambda_i(t_p) = (\hat{\lambda}_i^1(t_p) + \hat{\lambda}_i^2(t_p))/2$. We then average across phases (ignoring missing values) to obtain $\lambda_i(t) = \frac{1}{3}\sum_p \lambda_i(t_p)$. Any remaining missing values are imputed using the nearest-neighbor method.

We now have for each individual a vector λ_i with nine elements, one for each round from 2 to 10. We use the k-means algorithm to classify these in to separate clusters using the 'city-block' distance measure. This metric is preferred compared to the 'squared euclidean' metric, as the data is already in the same scale, but we do expect to have different variability across rounds. Using a non-linear distance metric would then put more weight on specific rounds. We repeat the clustering 1000 times with different random initializations and select the centroids that give the lowest within-cluster sums of data to centroid distances. To determine the number of clusters, we perform this exercise for different numbers of clusters and record the performance in terms of sum of distances. A higher number of clusters will always improve performance, but we observe that these gains are not substantial for more than 4 clusters (see Figure 13).

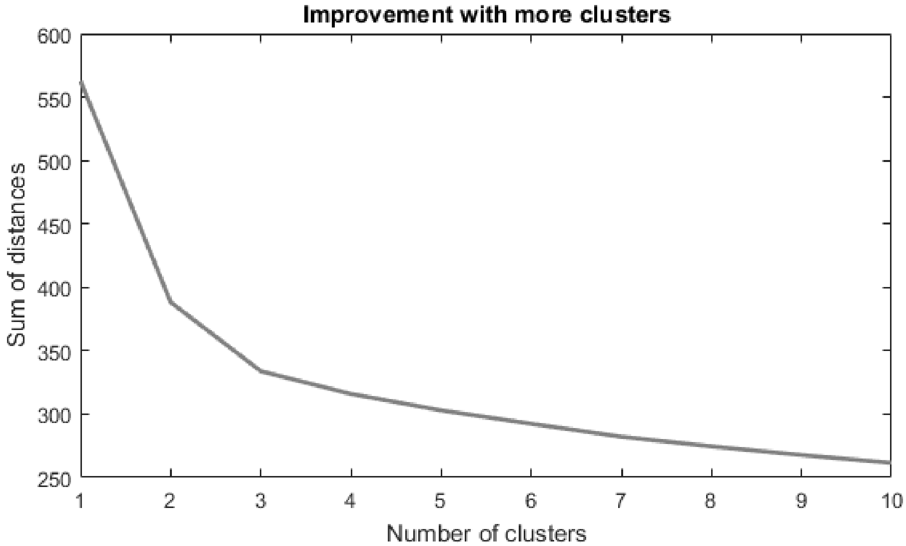


Figure 13: The graphs show the performance of the k-means clustering for different numbers of clusters. Lower values represent better performance.

To obtain the smooth decreasing functions showed in Figure 11, we fit a two-term power series model to each one of the four centroids. This can be viewed as a form of regularization that avoids overfitting the behavioral types to our data.

C Experimental Instructions

We present the instructions for the two fixed treatments F1 and F2. The relevant changes for the random treatment appear in footnotes.

INSTRUCTIONS

Thank you for participating in this session. Please remain quiet! You will be using the computer terminal for the entire experiment, and your decisions will be made via your computer terminals. Please DO NOT talk or make any other audible noises during the experiment. The use of mobile phones or other devices is prohibited. You are free to use the calculator provided. If you have any questions, raise your hand and your question will be answered so that everyone can hear.

General Instructions:

The experiment will take place in three parts. The remaining instructions refer to Part 1 of the experiment. Once part 1 is over you will be given instructions for Parts 2 and 3.

The experiment will involve a series of guesses. Each of you may earn different amounts. You also receive a £3 participation fee. Upon completion of the experiment, you will be paid individually and privately in room B33 upon presentation of the computer number you were assigned.

IMPORTANT:

The amount each participant earns, in today's experiment, depends only on his/her decisions and not on the decisions of other participants.

There is no specific time limit for making each guess. In order to finish the experiment on time we ask you to enter your guess in a reasonable amount of time. If a notification asking you to enter a guess appears on your screen please do so as soon as possible.

Part 1

The Task: In a tank there are 100,000 balls. These balls are either RED or BLUE. The number of balls of each colour is random and any combination is equally likely. You are asked to guess the number of RED balls in the tank. This number could be anywhere between 0 and 100,000.

Before making your guess, you observe a sample of 100 balls picked randomly from the tank. That is, the computer will inform you how many of the 100 balls in your sample were RED. You are then asked to enter a guess concerning the total number of red balls (between 0 and 100,000).

You will repeat this task three times. In other words you will make three guesses for three different tanks (filled with a different number of red and blue balls) and using a <u>different</u> sample each time. After each guess you will be given the correct number of RED balls and the points earned calculated as follows:

Points: For each guess you earn a maximum of 100 points if your guess is correct, while you earn fewer points for guessing wrong. The higher the difference between your guess and the correct number, the less points you earn. The exact number of points you earn in each round is given by the following formula:

$$\text{Points} = 100 - \text{Error Factor} \times \left(\frac{\text{Correct} - \text{Guess}}{1{,}000}\right)^2$$

If the result from the formula is negative, you earn zero points. You will be shown on your screen the exact value of the *Error Factor* for each guess you make.

Example

Suppose the number of red balls in the tank is 57,345. The following table illustrates examples of different guesses and the resulting number of points you would earn for different error factors.

Guess	**69,345**	**52,345**	**52,345**	**59,545**	**55,145**	**55,545**
Difference from correct	12,000	5,000	5,000	2,200	2,200	1,800
Error factor	1	1	10	10	25	25
Formula result	-44	75	-150	51.6	-21	19
Points	0	75	0	51.6	0	19

Given this formula, you maximize the expected number of points you earn in each round by making a guess that is as close as possible to your true estimate of the correct number of Red balls in the tank.

Out of the three guesses you will make in part 1, one will be selected randomly by the computer at the end of the experiment. The points you earned in the randomly selected guess will be transformed into monetary earnings. The exchange rate used is **£1 for every 85 points**. Notice that since all guesses can be chosen with the same probability, you cannot know for which of the guesses you will be paid. Therefore you should treat all guesses the same and make a guess as if you are going to be paid for it.

Part 2:

You will be assigned to a group that has 5 members. These groups are formed randomly and anonymously. You will interact exclusively within each group without knowing the identity of the other group members.

The Task: The task in this part has 5 rounds. Like in part 1, in a tank there are 100,000 balls. These balls are either RED or BLUE. The number of balls of each colour is random and any combination is equally likely. Each group member will be asked to guess the number of RED balls in the tank in each of the 5 rounds. This number could be anywhere between 0 and 100,000 and remains the same for all 5 rounds.

The task proceeds as follows: Before making a first guess, **each member observes a different sample of 100 balls picked randomly from the tank.**

Round 1: On your screen you will see the amount of RED balls in your sample of 100.

You are asked to make a guess about the number of red balls in the tank, as in Part 1.

Round 2: On your screen you will see the guess you made in round 1, as well as the guess(es) made in round 1 by some of the other group members. **You may observe a subset of one, two, or three other members. Each group member observes the guess(es) of a different subset of group members.** Furthermore, the fact that you observe a group member X does not necessarily mean that X observes you.[33]

You are asked to make a new guess about the number of RED balls in the tank.

Rounds 3-5: These rounds are the same as round 2. You see the guess(es) made previously by the group members you observe, and are asked to make a new guess.[34]

Payoffs: Again, you can earn a maximum of 100 points in each round if your guess is correct, while you earn fewer points for guessing wrong. The higher the difference between your guess and the correct number, the less points you earn. The exact number of points you earn in each round is given by the same formula as before:

$$\text{Points} = 100 - \text{Error Factor}_{round} \times \left(\frac{\text{Correct} - \text{Guess}}{1,000}\right)^2$$

Now the *Error Factor* is different in each round:

Round	1	2	3	4	5
Error Factor	1	5	10	20	25

If in a round the result from the formula is negative, you earn zero points in that round. After the 5 rounds are over, you will be shown how many points you earned in each round. At the end of the experiment the computer will randomly choose 1 out of the 5 rounds. The points you earned in this randomly chosen round will be transformed in to monetary earnings. The exchange rate used is **£1 for every 85 points**. Notice that since all 5 rounds can be chosen with the same probability, you cannot know for which of the rounds you will be

[33]**Round 2:** On your screen you will see the guess you made in round 1, as well as the guess(es) made in round 1 by some other randomly chosen group members. **You may observe one, two, or three other members.** The fact that you observe a group member X does not necessarily mean that X observes you.

[34]**Rounds 3-5:** These rounds are the same as round 2. Some group members are chosen randomly, you see their guess(es) from the previous round, and are asked to make a new guess.

paid. Therefore you should treat all rounds the same and make a guess as if you are going to be paid for it.

Remember:

1. **You will play 5 rounds. Your group and the members whose guesses you observe remain fixed during the whole time.**[35]
2. **There is a single tank and everybody in the group is guessing the number of RED balls in this tank. Each member observes a different sample of 100 balls and a different sample of group members.**
3. **Given the formula, you maximize the expected number of points you earn in each round by making a guess that is as close as possible to your true estimate of the correct number of Red balls in the tank.**
4. **The points you earn depend only on your guess and not on the guesses of other members.**

[35] **You will play 5 rounds. Your group and the number of members whose guesses you observe remain fixed during the whole time. The group members you observe are chosen randomly in each round.**

Part 3:

The composition of each group remains unchanged throughout all the experiment (same as in part 2). Remember that each group member observes the guess(es) of a different subset of group members, some will observe one, some two, and some three other group members.[36]

In this part there are 3 Phases of 10 Rounds each.

The Task: Now there are two tanks filled with 100,000 balls each. Tank 1 contains RED and BLUE balls, while Tank 2 contains GREEN and PURPLE balls. The number of balls of each colour in each tank is random and any combination is equally likely. You are asked to guess the correct number of RED balls in Tank 1 and of GREEN balls in Tank 2. **The number of RED balls in Tank 1 is not related to the number of GREEN balls in Tank 2.** These two numbers could be anywhere between 0 and 100,000.

As in part 2, before making a first guess, each participant observes 2 samples of 100 balls picked randomly: one sample for Tank 1, and one sample for Tank 2. **Remember that each participant observes different random samples.** Each phase then proceeds as follows:

Round 1: On your screen you will see the amount of RED balls in your sample from Tank 1 and the number of GREEN balls in your sample from Tank 2. You are asked to make a guess about the correct number of balls of the corresponding colour in each tank.

Round 2: As in part 2, on your screen you will see your guesses for each tank from round 1. You will also see the guesses made for each tank by the group members you observe from your group. After seeing their guesses you are asked to make new guesses about the number of RED balls in Tank 1 and GREEN balls in Tank 2.[37]

Rounds 3-10: As before, these rounds are the same as round 2. You see the guesses made in the previous rounds by the group members you observe, and make new guesses.[38]

History: Starting from the 2^{nd} phase in this part, you will have the opportunity to see the history of all guesses you and the group members you observe have made in previous phases. To access the history you just have to press the button on the top of your screen.

Payoffs: As in Parts 1 and 2, your payoff in each round is determined using the same formula. The guess you make for each tank enters the formula along with the correct number of balls of the corresponding colour. The Error Factor in each round is shown in the table below.

Round	1	2	3	4	5	6	7	8	9	10
Error Factor	1	5	10	15	15	15	20	20	20	25

The points you earn from each tank (maximum 100) are added together to give the total number of points for the round. At the end of each phase you will be shown how many points you earned in each round and from each tank. At the end of the experiment the computer

[36] **The composition of each group remains unchanged throughout all the experiment (same as in part 2). Remember that each group member observes the guess(es) of one, two, or three other group members in each round.**

[37] **Round 2:** As in part 2, on your screen you will see your guesses for each tank from round 1. You will also see the guesses made for each tank in the previous round by some other group members that are chosen randomly. After seeing their guesses you are asked to make new guesses about the number of RED balls in Tank 1 and GREEN balls in Tank 2.

[38] **Rounds 3-10:** As before, these rounds are the same as round 2. Some group members are chosen randomly, you see their guesses from the previous round, and are asked to make new guesses.

will randomly choose 1 out of the 10 rounds for each of the three phases. The points you earned in this randomly chosen round will be transformed into monetary earnings. The exchange rate used is **£1 for every 85 points**. Notice that since all 10 rounds of each phase can be chosen with the same probability, you cannot know for which of the rounds you will be paid. Therefore you should treat all rounds the same and make a guess as if you are going to be paid for it. Your monetary earning from each of these 3 rounds (one for each phase) will be added to your earnings from parts 1 and 2 and the show-up fee of £3. A screen will inform you about your total monetary earnings at the end of the experiment.

Remember:

1. **You will play 3 phases of 10 rounds. The groups and the members whose guesses you observe remain the same. What changes in each phase is the amount of balls in each tank and the samples observed by each member before making the first guess.**[39]
2. **In each phase there is a different amount of balls in each tank. The samples observed by each member before making the first guess are also different for each phase.**
3. **You maximize the expected number of points you earn in each round by making guesses that are as close as possible to your true estimates of the correct number of RED balls in Tank 1 and GREEN balls in Tank 2.**
4. **The points you earn depend only on your guess and not on the guesses of other members.**

[39] **You will play 3 phases of 10 rounds. The groups and their members remain the same. In each round some group members are chosen randomly and you observe their guesses in the previous rounds.**