Humanities & Social Sciences Communications

ARTICLE

https://doi.org/10.1057/s41599-024-03359-w OPEN

Convolution of individual and group identity: self-reliance increases polarisation in basic opinion model

Lennart Quante [1,2,6], Annika Stechemesser [1,3,4,6], Damian Hödtke [1,5] & Anders Levermann [1,4]✉

Opinion formation within society follows complex dynamics. Towards its understanding, axiomatic theory can complement data analysis. To this end, we propose an axiomatic model of opinion formation that aims to capture the interaction of individual conviction with social influence in a minimalist fashion. Despite only representing that (1) agents have an initial conviction with respect to a topic and are (2) influenced by their neighbours, the model shows the emergence of opinion clusters from an initially unstructured state. Here, we show that increasing individual self-reliance makes agents more likely to align their socially influenced opinion with their inner conviction which concomitantly leads to increased polarisation. The opinion drift observed with increasing self-reliance may be a plausible analogue of polarisation trends in the real-world. Modelling the basic traits of striving for individual versus group identity, we find a trade-off between individual fulfilment and societal cohesion. This finding from fundamental assumptions can serve as a building block to explain opinion polarisation.

[1] Potsdam Institute for Climate Impact Research (PIK), Member of the Leibniz Association, P.O. Box 6012 03, 14412 Potsdam, Germany. [2] Institute of Mathematics, University of Potsdam, Potsdam, Germany. [3] Mercator Research Institute on Global Commons and Climate Change (MCC), Berlin, Germany. [4] Institute of Physics, University of Potsdam, Potsdam, Germany. [5] Institute of Physics, Humboldt University of Berlin, Berlin, Germany. [6] These authors contributed equally: Lennart Quante, Annika Stechemesser. ✉email: anders.levermann@pik-potsdam.de

Introduction

Humans make thousands of decisions every day. Individual decisions are based on an opinion formation process that does not take place in isolation but evolves dynamically in relation to others. Although most micro-level decisions remain inconsequential, individual processes of opinion formation and mutual influence can in some cases have profound effects that affect society as a whole. A prominent example is elections, where each individual's vote, preceded by an opinion formation process, influences the outcome that determines governance. A second example is consumer decisions of individuals, which are often influenced by others (Grinblatt et al., 2008; Jansson et al., 2017) and can have the power to influence market trends. Understanding the mechanics of opinion formation, decision-making, and the mechanisms that underlie them is crucial for social processes at all scales. Experimental research on social influence dating back to Asch's research on conformity in the 1950s (Asch, 1956; Friend et al., 1990) has inspired empirical research into mechanisms of opinion polarisation (Moscovici and Zavalloni, 1969; Myers and Lamm, 1976). The need to further understand polarisation is highlighted by the observed increase in political polarisation in many countries (Abramowitz and McCoy, 2019; Geiger, 2014; Hohmann et al., 2023; Reiljan, 2020), the increased polarisation triggered by crises such as the Covid-19 pandemic (Charron et al., 2023; Druckman et al., 2021; Lobinska et al., 2022), and the effect of stronger polarisation on climate mitigation policies (Moore et al., 2022).

Opinion polarisation, understood here as a bi-modality in the opinion distribution describing the distance between different groups, has been the subject of a number of studies using agent-based modelling (ABM) that investigate the mechanisms and conditions that favour its emergence (Baldassarri and Page, 2021; Levin et al., 2021). Here, we add to this literature by using an ABM approach to investigate whether two opposing but fundamental human desires are plausible drivers of opinion polarisation: Belonging to a group and at the same time pursuing individual goals, i.e., to stand out from the group to some extent. The competition between these desires was for the first time proposed by Brewer in 1991 (Brewer, 1991), leading to the formulation of optimal distinctiveness theory (Leonardelli et al., 2010). It posits the need to balance assimilation and distinctiveness and provides the theoretical rationale for building our model. Similar approaches have been explored in ABMs proposed by Friedkin and Johnsen (Friedkin and Johnsen, 1990, 1997), systematic studies of conditions for polarisation (Hegselmann and Krause, 2002) as well as by Mäs et al. (Mäs and Flache, 2013; Mäs et al., 2010) (see Background and literature), but our approach differs from the literature in a number of ways.

First, the opinion evolution process in our model does not depend on homophily or bounded confidence assumptions, which would imply that agents are more willing to accept opinions similar to their own. Instead, the willingness of agents to be influenced by their neighbours' opinions, which may be similar or different to their own, is determined only by the agents' individual self-reliance, a characteristic describing their need for belonging/independence and their own intrinsic conviction (see "Methods" section for details). This implies that in our model, the polarisation patterns emerge naturally from the social dynamics. Second, our setup allows not only to analyse the continuous opinion distribution after model convergence, but agents also make a final, binary decision on a fictional issue, mimicking processes that end in a clear choice, such as voting or purchasing decisions. Finally, we do not only model polarisation patterns but also provide a measure of social cohesion based on the model result. Social cohesion has been defined along three core dimensions (Schiefer and van der Noll, 2017): quality of social relations, identification with society, and orientation towards the common good. To translate this into a simple but effective measure, we compute the opinion spread which measures the distance of opinions between the radical ends of the final opinion spectrum and is a proxy for the identification with the society represented by the average opinion as well as the ability of society to agree on the common good. By identifying a trade-off between social cohesion and the alignment of the agents' final decision with their intrinsic goal, we are able to show that the push and pull between individuality and cohesion holds not only at the individual level but also at the societal level.

The remainder of this paper is structured as follows: The background and literature section provides a brief overview of the literature on ABMs that model mechanisms of polarisation and explains how our model differs from this literature. A mathematical description of the model and its properties is given in the methods section. In the results section, we first show the emergence of divided groups from a random model initialisation (Fig. 2). Next, we explore how the share of agents with high self-reliance changes the strength of polarisation between the groups (Figs. 3 and 4). Finally, we compute the alignment between the intrinsic attitude of the agents and the final decision and show the trade-off between this decision alignment and societal cohesion (Figs. 5 and 6). In the discussion, we provide an interpretation of our results in a real-world context, discuss potential mechanisms, outline the limitations of the approach and conclude the paper.

Background and literature

The combination of complex systems theory with social sciences (Levin et al., 2021) is well-suited to explore mechanisms for polarisation in society (Baldassarri and Page, 2021). Agent-based models are a well-established tool to explore dynamic opinion formation and collective decision-making (Axelrod, 1997; Bianchi and Squazzoni, 2015; Sakoda, 1971; Schelling, 1971; Sobkowicz, 2020), and to investigate mechanisms of polarisation. Here, we give a brief overview of a few key mechanisms that have been explored in ABMs and how they differ from the setup in the model presented in this study.

A key mechanism for opinion formation that has been considered in ABMs is homophily, meaning the tendency to group with others who share traits, features or opinions close to our own (DeGroot, 1974). For example, homophily has been shown to amplify affinities between lifestyle and ideology (DellaPosta et al., 2015). Further, it can explain the formation of social groups and their size distribution (Korbel et al., 2023) or explain the emergence of social structures (Pham et al., 2022). Interdisciplinary studies have extended homophily-based models to study polarisation in a number of ways. For example, Dandekar et al. (2013) introduce the concept of biased assimilation, where individuals stick to their inherent belief if presented with inconclusive information on a complex issue, which leads to more extreme opinions. While the agents in our model are also influenced by their initial conviction or belief, we do not explore homophily combined with biased assimilation but self-reliance as a driver of the social dynamics, essentially exploring a different underlying mechanism. Our model also shares some similarities with the model proposed by Mäs et al. (2010) which uses individualisation introduced via an adaptive white-noise term in a homophily-based model. However, the models differ in two crucial points: First, the individualisation process in Mäs et al. (2010) is adaptive, meaning the need for individualisation is larger when the neighbourhood is more uniform, which is crucial to enable the clustering. In our model, individuality is parameterised by a constant self-reliance parameter that describes the inherent level of independence of an agent and is not changed in the opinion evolution process. Nevertheless, the push-and-pull of self-reliance and group-belonging driving our model leads to cluster emergence. In addition, the social dynamics in our model are not driven by an underlying homophily-process in contrast to Mäs et al. (2010). This is also a key difference between our model and the homophily-based model proposed in Mäs and Flache (2013), which shows that bi-polarisation can occur without individuals seeking to amplify differences with disliked others and instead can only be driven by an exchange of arguments in which arguments close to an individual's opinion have a stronger effect.

A further group of ABMs explores mechanisms that, in addition to assimilation or homophily effects, assume a repulsion effect, meaning that agents seek to distance themselves from others with dissimilar opinions (Baldassarri and Bearman, 2007; Flache and Macy, 2011; Mark, 2003; Martins et al., 2010). Prominent recent examples include Axelrod et al. (2021), which utilises this attraction-repulsion mechanism to explore the effects of polarisation on tolerance and responsiveness to other views, or Macy et al. (2021) which uses an attraction-repulsion model to illustrate asymmetric hysteresis trajectories in polarisation. Leonard et al. (2021) analyse the mechanisms of polarisation of elites (Kawakatsu et al., 2021; Kozlowski and Murphy, 2021) and consider the emergence of political factions under increasingly partisan identities and Chu et al. (2021) investigate the effects of political shocks on polarisation.

Finally, complex systems studies of specific conditions for polarisation in agent-based models explore the effects of coupled layers as a model of echo chambers (Gajewski et al., 2022) or combine polarisation and network evolution (Liu et al., 2023). Recent empirical studies on social networks are divided between finding polarising tendencies (Haroon et al., 2023) using YouTube and contrasting publications finding no increase in polarisation using data from Facebook (Guess et al., 2023; Nyhan et al., 2023). Other applications include the study of phenomena in financial markets and business, e.g., fraud in bitcoin markets (Fratrič et al., 2022; Zha et al., 2021).

Our model adds to the existing body of literature by presenting a reduced form ABM that focuses on the interplay between two key factors: the inclination to conform to the opinions of social contacts and the desire to maintain one's intrinsic beliefs. Rather than assuming specific assimilation or repulsion dynamics among agents, we focus on a timeless and inherent quality of individuals: their self-reliance, which reflects their quest for individual expression. Our results suggest that the push-and-pull between self-reliance and group affinity can explain the emergence of polarisation dynamics.

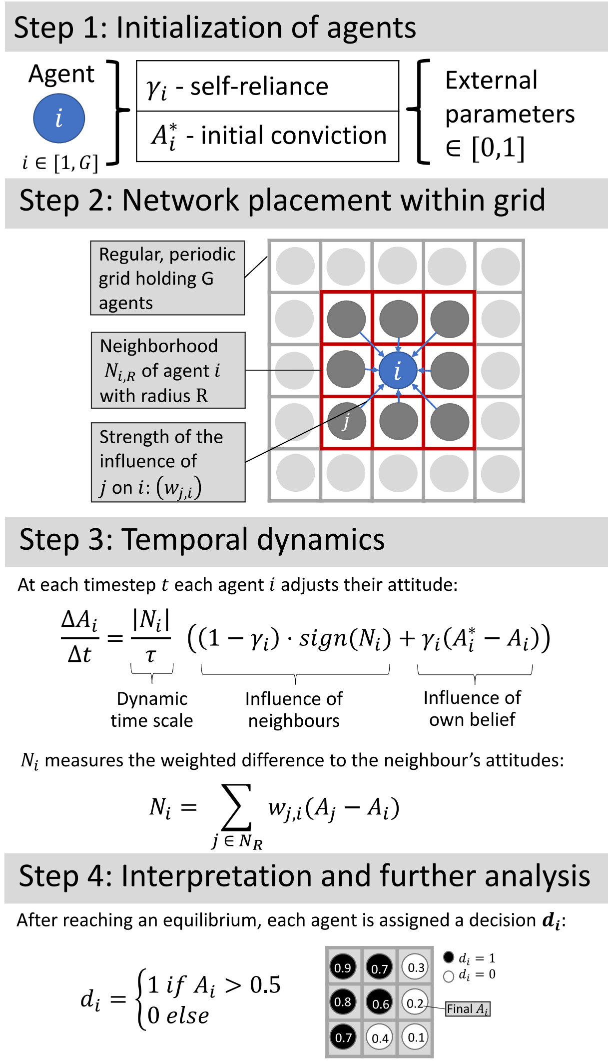

Fig. 1 Model setup and dynamics. Agents are initialised with a degree of self-reliance γ_i, expressing individualism, and an initial conviction A_i^*, which is also the starting value for attitude A_i. They are placed on a regular periodic grid with G agents—by default of size 100×100 yielding $G = 10.000$ agents—and assigned neighbourhoods based on Chebyshev distance. In the default specification, we use a Chebyshev distance of 1, i.e., a Moore neighbourhood with radius 1. At each time step, agents form an attitude A_i with the dynamics factoring in the influence of the neighbourhood and the strive towards the agent's initial conviction, both of which are influenced by self-reliance. Once the equilibrium is reached, each agent is assigned a final decision. Both the final opinion and the final decision are used for further analyses.

Methods

In this ABM, we explore opinion polarisation under the opposing forces of individuality and group-belonging. We formalise individuality (Fig. 1) by assigning every agent a continuous "self-reliance" parameter γ which describes how dependent on others the agent is in their opinion formation (continuous value between zero—"very dependent" to one "very self-reliant"). Further, every agent has an initial conviction A_i^*, which represents their intrinsic opinion on a topic (scaled continuously from zero—"full opposition" to one—"full agreement"). Both parameters are distributed uniformly in the basic version of the model. The agents are then randomly placed on a regular, periodic grid with G agents (100×100 agents by default, $G = 10{,}000$). All qualitative results are obtained by averaging over model ensembles of 100 varying initial distributions. Every agent is equally influenced by the eight neighbours around them, i.e., by their Moore neighbourhood with radius one. At each time step, every agent updates their attitude A based on the self-reliance weighted influence of their neighbours and the disparity between their own opinion and the opinions in their neighbourhood (see Fig. 1). Once the model has completed the final time step and reached a stable state, every agent makes a final decision for zero or one, determined by a specified threshold of their final attitude (default is 0.5). While simple, this model setup has multiple advantages. First, it is easily adjustable as the agent number, size, and weighting of neighbourhood influence and parameter distributions are modular. Second, it relies on very few input parameters, the self-reliance γ and initial conviction A_i^*. The genesis of opinions (going from initial conviction to final attitude) is not influenced by externally specified thresholds and all polarisation observed is emergent.

Finally, the interpretation is comprehensible, allowing an unobstructed view on the mechanism of self-reliance.

Technical model description. Here we give the technical details of the model described above. The model structure and dynamics are visualised in Fig. 1. We assume that each agent has two time-invariant attributes: An *initial conviction* $A_i^* \in [0,1]$ and a degree of *self-reliance* $\gamma_i \in [0,1]$, which expresses the agent's need for individualism with $\gamma_i = 1$ corresponding to individualistic opinion formation and $\gamma_i = 0$ implying opinion formation purely driven by the social neighbourhood. Furthermore, each agent's current opinion is described by a time-varying *attitude* parameter $A_i \in [0,1]$, which evolves as described by eq. (1). To initialise agent i at $t = 0$, the attitude is initialised equal to the initial conviction $A_i^* = A_i(0)$. A_i^* and γ_i are drawn from independent, random distributions. As a variation of the model, we provide additional results where the initial conviction A_i^* and the initial attitude $A_i(0)$ are independent of each other (Supplementary Figs. 22–25).

Agent i interacts with a set of neighbours M_i. In our simple example on a grid with periodic boundaries, i.e., a torus, all agents j with Chebyshev distance $d_C := \max(x_j - x_i, y_j - y_i) \le 1$ to agent i are part of the neighbourhood M_i. This is also known as a Moore neighbourhood with radius 1. The strength of the influence of j on i is given by the weight $w_{j,i}$, which we assume to be constant as $w_{j,i} := \frac{1}{\#M_i} \forall\, i, j$ for the sake of simplicity. For applications of our modelling framework on a network structure, these assumptions can be relaxed in the open-source model implementation. For robustness, we show results also for neighbourhoods up to $d_C \le 2$ in Supplementary Figs. 30–32 and find qualitatively robust results with only small variations.

A toy example for the agent evolution is visualised in Supplementary Fig. 1. At each time step t, agent i adjusts their *attitude* A_i according to the difference to the average attitude of their neighbourhood (defined via the *neighbourhood influence*, N_i) weighted by $1 - \gamma_i$ and according to their initial conviction A_i^* weighted by $\gamma_i |N_i|$. To avoid noise from sequential updating of the agents' attitudes, all agents update their attitude synchronously at time t based on their own attitude and their neighbours' attitudes of the previous time step $t - 1$. The self-reliance term is proportional to the attitude difference in the neighbourhood. Thus the more your opinion differs from that of your neighbours, the greater the influence of your initial conviction. Including a time scale τ, the dynamics of A_i are given by

$$\begin{aligned} \frac{\Delta A_i}{\Delta t} &= \frac{|N_i(t)|}{\tau}\left((1-\gamma_i)\frac{N_i(t)}{|N_i(t)|} + \gamma_i\left(A_i^* - A_i(t-1)\right)\right) \\ &= \frac{|N_i(t)|}{\tau}\left((1-\gamma_i)\mathrm{sign}\left(N_i(t)\right) + \gamma_i\left(A_i^* - A_i(t-1)\right)\right) \end{aligned} \tag{1}$$

with *neighbourhood influence* being defined as $N_i(t) = \sum_{j \in M_i} w_{j,i}(A_j(t-1) - A_i(t-1))$.

The opinion difference in the neighbourhood changes the time scale of the dynamics. The core of the dynamics (within the parenthesis) is a competition between the influence of the individual initial conviction and the neighbourhood which enters as a direction of change, since only the sign of the difference enters the dynamics.

The model is run until it converges to a steady state. Once the model has reached an equilibrium, each agent is assigned a final decision d_i which can be used in cases that require a binary decision:

$$d_i = \mathbf{1}_{A_i > 0.5} \tag{2}$$

The derivation of equilibrium conditions is given in the supplementary material—section "Equilibrium conditions"—and yields two qualitatively different equilibria:

$$N_i = 0 \text{ and } A_i^* - A_i = \frac{1-\gamma_i}{\gamma_i}\mathrm{sign}\left(N_i\right). \tag{3}$$

Measures of self-fulfilment and social cohesion

Decision alignment as a measure of self-fulfilment. Here we understand self-fulfilment as the extent to which an agent is following their personal initial conviction. We formalise this by considering the alignment of initial conviction and the final decision of an agent which is taken after being exposed to the neighbourhood influences. To this end, we define the initial decision $\tilde{d}_i$ of agent i as $\tilde{d}_i := \mathbf{1}_{A_i^* > 0.5}$, where A_i^* is the initial conviction. Recall that the final decision d_i is defined as $d_i = \mathbf{1}_{A_i > 0.5}$. Then the decision alignment is posed as

$$\delta_i := \mathbf{1}_{d_i = \tilde{d}_i}. \tag{4}$$

Thus, a positive decision alignment is achieved if the initial conviction and the final decision lead an agent to the same binary choice. The societal level of self-fulfilment is computed as the average decision alignment, formally expressed as

$$\Delta := \frac{\sum_i \delta_i}{G}, \tag{5}$$

where G is the total number of agents. In Fig. 5 we show how the societal level of self-fulfilment changes for different levels of average self-reliance.

In addition to these empirical results, we can also analytically derive the dependence of the average decision alignment Δ on the distribution of self-reliance γ_i using a mean-field approximation approach. This means we are deriving average results for society using the simplifying assumption of independence of A_i and N_i, which is only true when considering an average across society as a whole, but not for individual agents. Specifically, we approximate the expected decision alignment of agent i with self-reliance γ_i based on the possible equilibria identified previously eq. (3). The expectation of δ_i in the final equilibrium state can be estimated as

$$\begin{aligned} \Delta = \mathbb{E}\left[\delta_i\right] &= \left[\mathbb{P}(N_i < 0) + \mathbb{P}(N_i > 0)\right]\left[\tfrac{1}{2} + \mathbb{P}\left(\mathcal{U}(0, 0.5) > \tfrac{1-\gamma_i}{\gamma_i}\right)\right] + \tfrac{1}{2}\mathbb{P}(N_i = 0) \\ &= \tfrac{1}{2} + \left[\mathbb{P}(N_i \neq 0)\right]\mathbb{P}\left(\mathcal{U}(0, 0.5) > \tfrac{1-\gamma_i}{\gamma_i}\right) \end{aligned} \tag{6}$$

Considering $\mathbb{P}\left(\mathcal{U}(0, 0.5) > \frac{1-\gamma_i}{\gamma_i}\right)$, it follows that only individuals with $\gamma_i \ge \frac{2}{3}$ can contribute to this term, since

$$0.5 \ge \frac{1-\gamma_i}{\gamma_i} \iff 0.5\gamma_i \ge 1 - \gamma_i \iff \gamma_i \ge \frac{2}{3}$$

This approximation of average decision alignment allows for a comparison of our simulation results with analytical expectations, as we present in Fig. 5. We provide a detailed derivation of this expected average decision alignment (eq. (6)) in the supplementary material (Detailed derivation of eq. (6)).

Opinion spread as a measure of social cohesion. To measure social cohesion, we consider the opinion spread ω between the extreme ends of society as the difference of 90th and 10th percentile of the final attitude, i.e., for $\mathbb{A}_{final}$ denoting the final attitudes of all agents we obtain:

$$\omega := Pct_{90}(\mathbb{A}_{final}) - Pct_{10}(\mathbb{A}_{final}). \tag{7}$$

Thus, if the opinion spread is large, the distance between the radical ends of the opinion spectrum is large, implying that the identification with the society represented by an average opinion is low. In addition, a large opinion spread may pose obstacles

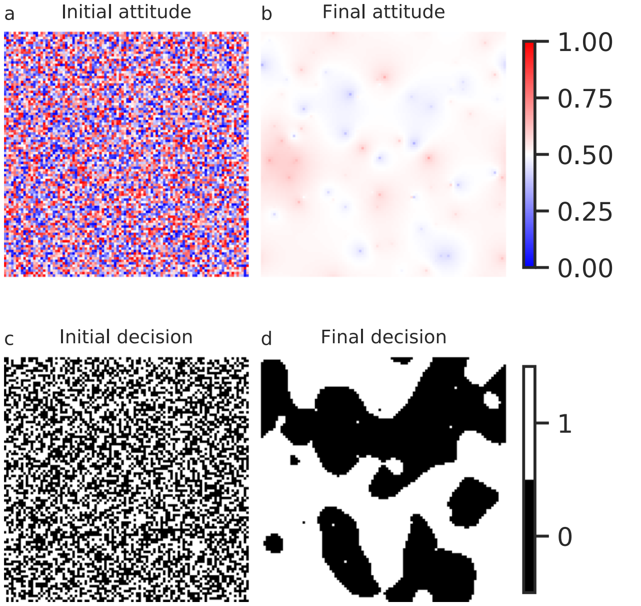

Fig. 2 Trade-off between belonging and individualism leads to the emergence of stable, opposing opinion clusters. a shows the uniformly distributed initial attitude for each agent. Agents are placed randomly on the periodic grid. **b** shows the final attitude after evolving the model for 1000 time steps. Opinion clusters emerge around a few agents with strong opinions and many agents with more moderate views. **c** visualises the binary initial decision, which is based on the initial conviction (threshold of 0.5). No clusters are visible. In contrast, the final decision (**d**), which is based on the final attitude, shows clear opinion clusters.

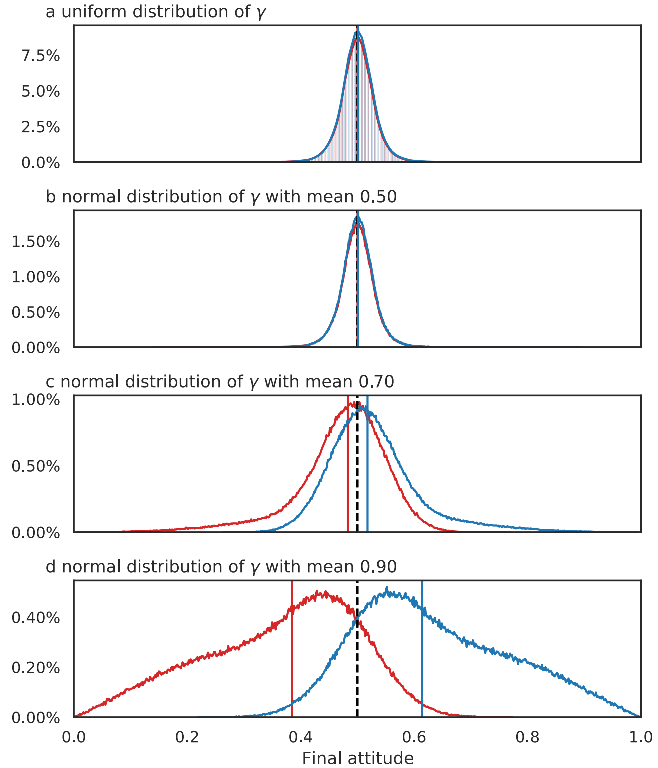

Fig. 3 Opinion spread increases in more self-reliant societies. Panels show histograms of the final attitude: **a** for uniformly parameterised population, **b** for normally distributed γ with a mean of 0.5, **c** for normally distributed γ with a mean of 0.7, and **d** for normally distributed γ with a mean of 0.9. Red shows initial convictions < 0.5, blue shows initial convictions ≥ 0.5. Solid lines show medians, the dashed black line marks 0.5.

when deciding on a shared beneficial solution. Therefore, we interpret a large opinion spread as an indication of low social cohesion whereas a smaller opinion spread points to higher social cohesion.

For robustness, we also show results for varied percentile thresholds of opinion spread in the supplementary information, as referenced in the results.

Results

Emergence of a stable and opinion-divided society. For uniform distributions of initial attitude and self-reliance, we consistently find an emergence of stable, polarised decision clusters. Figure 2 shows the difference between the initial attitude (a) and initial decision (c) as well as the final attitude (b) and decision (d) exemplary on a 100×100 grid. In the initial state, no patterns or structures are visible. After evolving the model for 1000 time steps, we see clusters of agents with final attitudes (b) that strongly diverge from the threshold (here 0.5) in either direction surrounded by agents with moderate opinions. This suggests that neighbourhoods can be overly influenced by few self-reliant agents, which mirrors patterns observed in the structure of scale-free social networks, where few agents are influential (many connections) and many agents are influenced (few connections). Visualising the final decision (d) shows multiple distinct decision clusters. For both possible decisions (one or zero) there is one large cluster and multiple smaller clusters.

We expand on the uniform initialisation by considering normal distributions of agent self-reliance with varying means (Supplementary Fig. 2) as well as varying standard deviations (Supplementary Figs. 7, 12, and 17). Initial attitude as well as agent placement remain as before. Again we observe the formation of stable, opposing opinion clusters in all realisations. The size and balance of clusters change with the distribution of individualistic agents. Smaller average self-reliance induces the formation of larger clusters, larger average self-reliance implies fine clustering. A small mean and standard deviation of self-reliance leads to smaller differences in final attitude (Supplementary Fig. 7), showing that variation of self-reliance is necessary for strong clustering to emerge. As an alternative to the uniform neighbourhoods, we initialise the grid placing agents with a placement probability $p = 0.75$ which leads to a non-uniform, partially populated grid. For this scenario, we observe similar clustering patterns (Supplementary Fig. 26). Varying the neighbourhood radius from 1 to 2 also conserves the general pattern of emerging clusters (Supplementary Fig. 30).

Overall, these results show that the proposed simple decision model leads to non-trivial decision patterns based on the trade-off between social influences and the reliance on personal attitude. Few, strong-minded agents are sufficient to create large, stable clusters.

Societal polarisation increases with more individualistic agents. The exemplary results in Fig. 2 suggest that opinion clusters form around a few agents with a very strong final attitude and many agents with a more moderate final attitude. We systematically explore this by varying the uniform distribution of initial variables in an ensemble of societies of 10,000 agents over 100 runs each. To evaluate how the attitude evolved, we split agents according to their initial conviction (threshold of 0.5, i.e., $A_i^* < 0.5$ and $A_i^* \geq 0.5$) and visualise the distribution of the final attitude separately for each group (Fig. 3a).

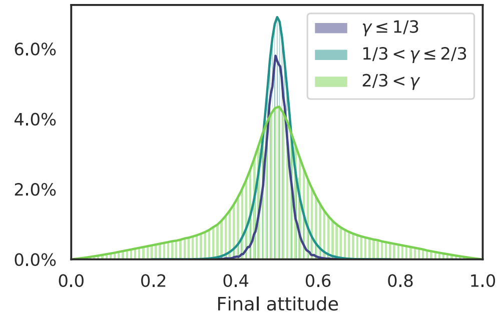

Fig. 4 Highly self-reliant individuals have more polarised opinions. Histograms of the final attitude for bins of individual self-reliance γ based on an ensemble of normal distributions of γ with means between 0.5 and 0.9.

We find that the final attitude is approximately normally distributed but with long tails. The majority of agents stick with their inherent decisions, but their attitudes become more moderate. In contrast to the initial uniform distribution, the median is much closer to 0.5 in the final distribution for both groups. A tail of agents with small or large attitudes remains. The joint near-normal distribution of final attitudes matches distributions of opinions observed in real life (Baldassarri and Gelman, 2008; Geiger, 2014; Hohmann et al., 2023). To investigate the impact of changing distributions of self-reliance, we consider normally distributed self-reliance (γ) instead of a uniform distribution. All other initial values remain the same. We fix the standard deviation as $\sigma(\gamma) = 0.1$ and consider different means of the normal distribution ($\mu(\gamma) = 0.5, \mu(\gamma) = 0.7, \mu(\gamma) = 0.9$). We then evaluate how the distribution of the final attitude changes in response. With an increasing mean of self-reliance, i.e., a higher share of individualistic agents, the variance of the final attitude distribution increases and the difference in the final mean attitude of the two groups grows (Fig. 4b–d). This shows an opinion drift towards stronger polarisation. To assess which agents populate the opinion tails, we merge the runs with varying means ($\mu \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$) into a shared ensemble and visualise the distribution of the final attitude separately for different thresholds of self-reliance ($\gamma \le \frac{1}{3}, \frac{1}{3} < \gamma \le \frac{2}{3}, \frac{2}{3} < \gamma$). As shown in Fig. 4, agents that have higher self-reliance populate the tails of the final attitude distribution, such that these highly self-reliant individuals drive the polarisation of society at large, an observation also made in previous studies of more complex models (Turner and Smaldino, 2018).

Varying the standard deviation of the normal distribution of self-reliance compared to Fig. 3 (Supplementary Figs. 8, 13, and 18) we find that the characteristics of resulting distributions persist, with lower standard deviations leading to slightly more concentrated attitudes (Supplementary Fig. 8) and higher standard deviation to a larger spread (Supplementary Fig. 18). This also holds for Fig. 4 as shown in Supplementary Figs. 9, 14, and 19. We also relax the assumption that the initial attitude A_i equals the inherent conviction A_i^*, observing a similar increase in opinion spread with increasing self-reliance (Supplementary Figs. 22–25). For a partially populated grid, these observations remain robust as well (Supplementary Figs. 27 and 28). For an increased neighbourhood radius (2 instead of 1), the general tendency of increasing opinion spread with increasing self-reliance remains, while the stronger connection among all agents leads to a slight tendency for a more central distribution Supplementary Fig. 31).

The opinion drift and polarisation observed here for increasing numbers of self-reliant agents tie in with empirical observations of societies that experience a growing number of citizens with strong, opposing political opinions over time. An example is the political polarisation of the United States, where opinion polls show near-normal distributions that drift apart over time in recent times (Geiger, 2014; Hohmann et al., 2023). This suggests that the mechanism we model has potential to map real-world phenomena.

Trade-off between self-fulfilment and social cohesion. Every agent is equipped with an initial conviction which represents their individual stance on a topic. Their final attitude then arises dynamically in the field of tension between the influence of their neighbours and their initial conviction, weighted according to self-reliance. We now assess if the decision the agent would have taken based on their initial conviction aligns with their final decision d_i taken based on the final attitude. To this end, we recall that the decision alignment indicates if the initial decision $\tilde{d}_i$ and final decision d_i match (see Methods, eq. (4)). Considering the average decision alignment Δ across all agents (see Methods, eq. (5)) in dependence of the mean self-reliance $\langle\gamma\rangle$ shows that societies with more self-reliant agents achieve a higher average decision alignment: the agents are more prone to follow their own beliefs independently of their neighbourhoods (Fig. 5a). Comparing these model results to an analytical approximation of this behaviour (grey line in Fig. 5a, see Methods equation eq. (6)) shows that the society-wide mean self-reliance might be used as a proxy to estimate results emerging from simulations based on individual agents.

Next, we show that the opinion spread, measuring the difference of the 90th and 10th percentile of the final attitude (see Methods, eq. (7)), also increases with self-reliance, indicating a wider spread of attitudes and thus a potential decrease in social cohesion as opinions in society drift apart (Fig. 5b). This means we observe a trade-off: If the number of self-reliant agents is low, the opinion spread is small and social cohesion is high. However, the societal decision alignment is also lower as agents have to compromise more. We visualise this trade-off in Fig. 6 plotting decision alignment against opinion spread.

Varying the percentile levels defining the opinion spread leads to qualitatively similar results compared to Fig. 5 (Supplementary Figs. 3 and 5) as well as for Fig. 6 (Supplementary Figs. 4 and 6). Varying the standard deviation of self-reliance does not change the main observations either, as shown in Fig. 5 in Supplementary Figs. 10, 15, and 20 and Fig. 6 in Supplementary Figs. 11, 16, and 21. Again, also for the partially populated grid a similar trade-off can be observed (Supplementary Fig. 29). For an increased neighbourhood radius, the trade-off between decision alignment and opinion spread with increasing self-reliance remains, while the stronger connection among all agents induces a slightly delayed, but subsequently steeper increase of the opinion spread (Supplementary Fig. 32).

Discussion

Examining underlying drivers for human decision-making is crucial for understanding societal processes. In this study, we use an agent-based approach to model opinion formation against the push-and-pull of individuality and group-belonging. This approach is motivated by optimal distinctiveness theory (Brewer, 1991) and the related concept of self-determination (Deci and Ryan, 2012). Even though there are individual studies exploring optimal distinctiveness theory in agent-based modelling (Smaldino et al., 2012), these studies do not discuss the broader implications for societal opinion formation. We pursue a strategy of model-driven exploration of behavioural patterns, positing that relatively simple rules for agents can reproduce emerging

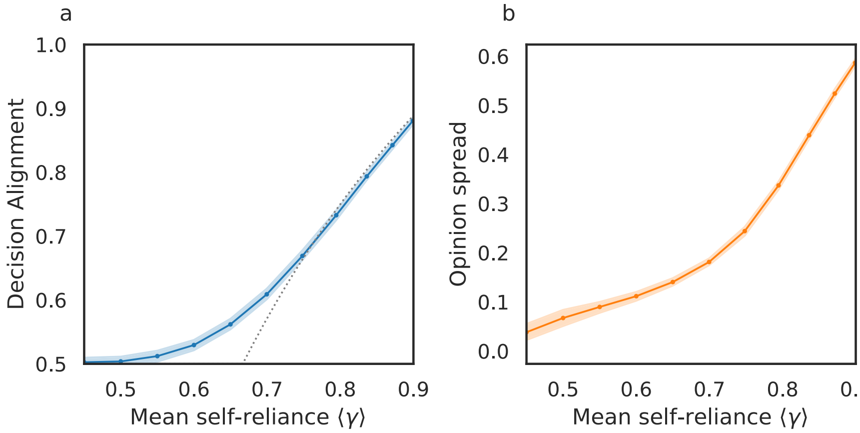

Fig. 5 Trade-off between decision alignment and social cohesion. a shows the decision alignment in dependence on the average number of self-reliant agents. The decision alignment is computed as the average of the differences between the agents' initial decision and their final decision after evolving the model. The grey dots show an analytical approximation based on a mean-field approximation (Assuming 25% of agents are in equilibrium with their neighbours, i.e., $\mathbb{P}(N_i = 0) = 0.25$). **b** shows the opinion spread in dependence on the average number of self-reliant agents measured as the difference between the 90th and 10th percentiles of the distribution of the final attitude. The initial opinion spread is about 0.8 due to the uniform distribution of the initial conviction. If the opinion spread is large, societal opinions are drifting apart and social cohesion lowers. Hence, there is a trade-off between higher personal decision alignment with more self-reliant agents and more social cohesion with less self-reliant agents. The average number of self-reliant agents corresponds to the mean of a normal distribution with mean γ and standard deviation $\sigma = 0.1$ from which the self-reliance was sampled. Confidence bands show the [5, 95] confidence interval based on 100 simulations with varying initial conditions.

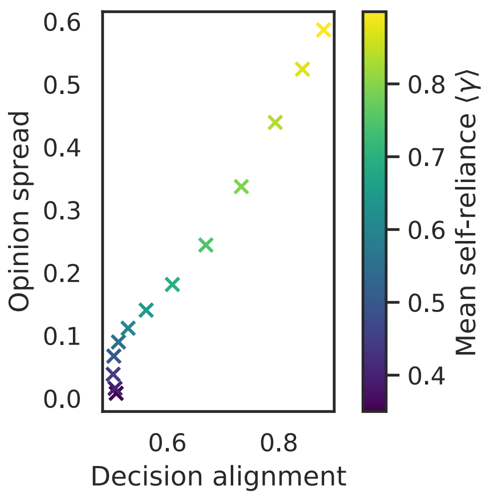

Fig. 6 Trade-off between opinion spread and alignment with an inherent decision. Decision alignment and opinion spread increase for higher mean self-reliance. Markers show mean values for colour-coded mean self-reliance.

phenomena of society as demonstrated in recent work on emerging ostracism (Lindström and Tobler, 2018).

We find that even in a simple model, these opinion dynamics lead to polarisation with stable, opposing opinion clusters. Further, with an increasing number of independent agents, which are difficult to influence, a stronger drift between opposing views emerges, culminating in a trade-off between the agent-level individual alignment with their personal opinion and societal cohesion.

Our findings align with empirical evidence for opinion drift that leads to stronger societal polarisation such as the political polarisation of the United States in the last decades (Geiger, 2014; Hohmann et al., 2023). The data show the drifting apart of the mean distributions of political alignment between Republican and Democratic voters resp. members of Congress which maps the drift in attitude means we observe in Fig. 3. This suggests that the juxtaposition between belonging and self-reliance might be a mechanism that promotes social polarisation which complements other polarisation mechanisms that have been examined in the theoretical and empirical literature. The fact that more self-reliant agents lead the opinions of others can be connected to recent empirical work on political opinions (Goldenberg et al., 2023), which shows that complementing homophily with the preference to agree with more radical options, named acrophily, might contribute to political segregation. Thus our finding of a trade-off between self-reliance and the higher frequency of more extreme opinions may be a harbinger of a divided opinion spectrum when self-reliance crosses a societal threshold. In contrast to theoretical work (Chuang et al., 2016) or models (Axelrod et al., 2021) exploring polarisation of society, we do not explicitly model repulsion from other opinions. Polarisation emerges from a stronger drive towards a personal inherent opinion. While this is a variation of similar mechanisms, the proposed interpretation as reliance on your own opinion aligns with psychological concepts like optimal distinction theory (Brewer, 1991) or self-determination (Deci and Ryan, 2012). Introducing a bias towards a personal opinion, our model is related to but distinct from the inclusion of biased assimilation (Dandekar et al., 2013). Even though our model is simple, it reproduces multiple previous findings of more complex models. In particular, we find polarisation emerging from a random initial state, the tendency to follow more radical agents and a connected trade-off between societal cohesion and individuality.

The minimal model presented here intentionally omits additional mechanisms. Instead of striving for a granular, general model, we present a new perspective on how polarisation may be driven by the push-and-pull of individuality and group-belonging. We have shown that by introducing an intuitive component of self-reliance into an averaging neighbours model, polarisation emerges from an increasing spread of opinions and clustering occurs. Thus, considering the individual reliance on inherent opinion complementing adjustment to opinions of social contacts may contribute to explaining polarisation in human decision-making.

Data availability

The simulation data that support the findings of this study are openly available at the public repository for this publication with the identifier https://doi.org/10.5281/zenodo.11640071.

Code availability

The model and analysis code are available open-source on Github as referenced at https://doi.org/10.5281/zenodo.11640071.

Received: 2 November 2023; Accepted: 17 June 2024;
Published online: 27 June 2024

References

Abramowitz A, McCoy J (2019) United States: racial resentment, negative partisanship, and polarization in Trump's America. Ann Am Acad Political Soc Sci 681:137–156
Asch SE (1956) Studies of independence and conformity: I. A minority of one against a unanimous majority. Psychol Monogr Gen Appl 70:1–70
Axelrod R (1997) The dissemination of culture: a model with local convergence and global polarization. J Confl Resolut 41:203–226
Axelrod R, Daymude JJ, Forrest S (2021) Preventing extreme polarization of political attitudes. Proc Natl Acad Sci 118:e2102139118
Baldassarri D, Bearman P (2007) Dynamics of political polarization. Am Sociol Rev 72:784–811
Baldassarri D, Gelman A (2008) Partisans without constraint: political polarization and trends in american public opinion. AJS; Am J Sociol 114:408–446
Baldassarri D, Page SE (2021) The emergence and perils of polarization. Proc Natl Acad Sci 118:e2116863118
Bianchi F, Squazzoni F (2015) Agent-based models in sociology. WIREs CompStat 7:284–306
Brewer MB (1991) The social self: on being the same and different at the same time. Personal Soc Psychol Bull 17:475–482
Charron N, Lapuente V, Rodríguez-Pose A (2023) Uncooperative society, uncooperative politics or both? Trust, polarization, populism and COVID-19 deaths across European regions. Eur J Political Res 62:781–805
Chu OJ, Donges JF, Robertson GB, Pop-Eleches G (2021) The microdynamics of spatial polarization: a model and an application to survey data from Ukraine. Proc Natl Acad Sci 118:e2104194118
Chuang Y-L, D'Orsogna M, Chou T (2016) A bistable belief dynamics model for radicalization within sectarian conflict. Q Appl Math 75:19–37
Dandekar P, Goel A, Lee DT (2013) Biased assimilation, homophily, and the dynamics of polarization. Proc Natl Acad Sci 110:5791–5796
Datseris G, Vahdati AR, DuBois TC (2022) Agents.jl: a performant and feature-full agent-based modeling software of minimal code complexity. Simulation 0:003754972110688
Deci EL, Ryan RM (2012) Self-determination theory. In: Handbook of Theories of Social Psychology. vol 1. SAGE Publications Ltd, London. pp. 416–437
DeGroot MH (1974) Reaching a consensus. J Am Stat Assoc 69:118–121
DellaPosta D, Shi Y, Macy M (2015) Why do liberals drink lattes? AJS Am J Sociol 120:1473–1511
Druckman JN, Klar S, Krupnikov Y, Levendusky M, Ryan JB (2021) Affective polarization, local contexts and public opinion in America. Nat Hum Behav 5:28–38
Flache A, Macy MW (2011) Small worlds and cultural polarization. J Math Sociol 35:146–176
Fratrič P, Sileno G, Klous S, van Engers T (2022) Manipulation of the Bitcoin market: an agent-based study. Financial Innov 8:60
Friedkin NE, Johnsen EC (1990) Social influence and opinions. J Math Sociol 15:193–206
Friedkin NE, Johnsen EC (1997) Social positions in influence networks. Soc Netw 19:209–222
Friend R, Rafferty Y, Bramel D (1990) A puzzling misinterpretation of the Asch 'conformity' study. Eur J Soc Psychol 20:29–44
Gajewski LG, Sienkiewicz J, Holyst JA (2022) Transitions between polarization and radicalization in a temporal bilayer echo-chamber model. Phys Rev E 105:024125
Geiger A (2014). Political polarization in the American Public, Pew Research Center
Goldenberg A, Abruzzo JM, Huang Z, Schöne J, Bailey D, Willer R, Halperin E, Gross JJ (2023) Homophily and acrophily as drivers of political segregation. Nat Hum Behav 7:219–230
Grinblatt M, Keloharju M, Ikäheimo S (2008) Social influence and consumption: evidence from the automobile purchases of neighbors. Rev Econ Stat 90:735–753
Guess AM, Malhotra N, Pan J, Barberá P, Allcott H, Brown T, Crespo-Tenorio A, Dimmery D, Freelon D, Gentzkow M, González-Bailón S, Kennedy E, Kim YM, Lazer D, Moehler D, Nyhan B, Rivera CV, Settle J, Thomas DR, Thorson E, Tromble R, Wilkins A, Wojcieszak M, Xiong B, de Jonge CK, Franco A, Mason W, Stroud NJ, Tucker JA (2023) How do social media feed algorithms affect attitudes and behavior in an election campaign? Science 381:398–404
Haroon M, Wojcieszak M, Chhabra A, Liu X, Mohapatra P, Shafiq Z (2023) Auditing YouTube's recommendation system for ideologically congenial, extreme, and problematic recommendations. Proc Natl Acad Sci 120:e2213020120
Hegselmann R, Krause U (2002) Opinion dynamics and bounded confidence models, analysis, and simulation. J Artif Soc Soc Simulation 5(3)-2:1–33
Hohmann M, Devriendt K, Coscia M (2023) Quantifying ideological polarization on a network using generalized Euclidean distance. Sci Adv 9:eabq2044
Jansson J, Pettersson T, Mannberg A, Brännlund R, Lindgren U (2017) Adoption of alternative fuel vehicles: Influence from neighbors, family and coworkers. Transp Res Part D Transp Environ 54:61–73
Kawakatsu M, Lelkes Y, Levin SA, Tarnita CE (2021) Interindividual cooperation mediated by partisanship complicates Madison's cure for "mischiefs of faction". Proc Natl Acad Sci 118:e2102148118
Korbel J, Lindner SD, Pham TM, Hanel R, Thurner S (2023) Homophily-based social group formation in a spin glass self-assembly framework. Phys Rev Lett 130:057401
Kozlowski AC, Murphy JP (2021) Issue alignment and partisanship in the American public: revisiting the 'partisans without constraint' thesis. Soc Sci Res 94:102498
Leonard NE, Lipsitz K, Bizyaeva A, Franci A, Lelkes Y (2021) The nonlinear feedback dynamics of asymmetric political polarization. Proc Natl Acad Sci 118:e2102149118
Leonardelli GJ, Pickett CL, Brewer MB (2010) Chapter 2—Optimal distinctiveness theory: a framework for social identity, social cognition, and intergroup relations. In: Zanna MP, Olson JM (eds) Advances in experimental social psychology. vol. 43. Academic Press. pp. 63–113
Levin SA, Milner HV, Perrings C (2021) The dynamics of political polarization. Proc Natl Acad Sci 118:e2116950118
Lindström B, Tobler PN (2018) Incidental ostracism emerges from simple learning mechanisms. Nat Hum Behav 2:405–414
Liu J, Huang S, Aden NM, Johnson NF, Song C (2023) Emergence of polarization in coevolving networks. Phys Rev Lett 130:037401
Lobinska G, Pauzner A, Traulsen A, Pilpel Y, Nowak MA (2022) Evolution of resistance to COVID-19 vaccination with dynamic social distancing. Nat Hum Behav 6:193–206
Macy MW, Ma M, Tabin DR, Gao J, Szymanski BK (2021) Polarization and tipping points. Proc Natl Acad Sci 118:e2102144118
Mark NP (2003) Culture and competition: homophily and distancing explanations for cultural niches. Am Sociol Rev 68:319–345
Martins TV, Pineda M, Toral R (2010) Mass media and repulsive interactions in continuous-opinion dynamics. Europhys Lett 91:48003
Moore FC, Lacasse K, Mach KJ, Shin YA, Gross LJ, Beckage B (2022) Determinants of emissions pathways in the coupled climate-social system. Nature 603:103–111
Moscovici S, Zavalloni M (1969) The group as a polarizer of attitudes. J Personal Soc Psychol 12:125–135
Myers DG, Lamm H (1976) The group polarization phenomenon. Psychol Bull 83:602–627
Mäs M, Flache A (2013) Differentiation without distancing. explaining bi-polarization of opinions without negative influence. PLoS ONE 8:e74516
Mäs M, Flache A, Helbing D (2010) Individualization as driving force of clustering phenomena in humans. PLoS Comput Biol 6:e1000959
Nyhan B, Settle J, Thorson E, Wojcieszak M, Barberá P, Chen AY, Allcott H, Brown T, Crespo-Tenorio A, Dimmery D, Freelon D, Gentzkow M, González-Bailón S, Guess AM, Kennedy E, Kim YM, Lazer D, Malhotra N, Moehler D, Pan J, Thomas DR, Tromble R, Rivera CV, Wilkins A, Xiong B, de Jonge CK, Franco A, Mason W, Stroud NJ, Tucker JA (2023) Like-minded sources on Facebook are prevalent but not polarizing. Nature 620:137–144
Pham TM, Korbel J, Hanel R, Thurner S (2022) Empirical social triad statistics can be explained with dyadic homophylic interactions. Proc Natl Acad Sci 119:e2121103119
Reiljan A (2020) 'Fear and loathing across party lines' (also) in Europe: Affective polarisation in European party systems. Eur J Political Res 59:376–396
Sakoda JM (1971) The checkerboard model of social interaction. J Math Sociol 1:119–132
Schelling TC (1971) Dynamic models of segregation. J Math Sociol 1:143–186
Schiefer D, van der Noll J (2017) The essentials of social cohesion: a literature review. Soc Indic Res 132:579–603

Smaldino P, Pickett C, Sherman J, Schank J (2012) An agent-based model of social identity dynamics. J Artif Soc Soc Simul 15:7
Sobkowicz P (2020) Whither now, opinion modelers? Front Phys 8:587009
Turner MA, Smaldino PE (2018) Paths to polarization: how extreme views, miscommunication, and random chance drive opinion dynamics. Complexity 2018:e2740959
Zha Q, Kou G, Zhang H, Liang H, Chen X, Li C-C, Dong Y (2021) Opinion dynamics in finance and business: a literature review and research opportunities. Financ Innov 6:44

Acknowledgements
We thank the developers of the Agents.jl ((Datseris et al., 2022) package for providing the basis for our model development and Maximilian Kotz for helpful discussions. This research has received funding from the German Federal Ministry of Education and Research (BMBF) under the research projects QUIDIC (01LP1907A), from the Volkswagen Foundation and the Werner Siemens Foundation. The authors gratefully acknowledge the European Regional Development Fund (ERDF), the German Federal Ministry of Education and Research, and the Land Brandenburg for supporting this project by providing resources on the high-performance computer system at the Potsdam Institute for Climate Impact Research.

Author contributions
LQ, AS, and AL designed the study. LQ and AS developed the model code. DH conducted an exploratory analysis and extended the model following discussions with all authors. DH analysed the supplementary two-agent case. LQ and AS wrote the manuscript. All authors discussed the results and approved the final manuscript.

Funding
Open Access funding enabled and organized by Projekt DEAL.

Competing interests
The authors declare no competing interests.

Ethical approval
This article does not contain any studies with human participants performed by any of the authors.

Informed consent
This article does not contain any studies with human participants performed by any of the authors.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-024-03359-w.

Correspondence and requests for materials should be addressed to Anders Levermann.