nature human behaviour

Article

https://doi.org/10.1038/s41562-023-01715-5

Resampling reduces bias amplification in experimental social networks

Received: 21 September 2022

Accepted: 6 September 2023

Published online: 16 October 2023

Check for updates

Mathew D. Hardy [1,5] ✉, Bill D. Thompson [2,5], P. M. Krafft [3] & Thomas L. Griffiths [1,4]

Large-scale social networks are thought to contribute to polarization by amplifying people's biases. However, the complexity of these technologies makes it difficult to identify the mechanisms responsible and evaluate mitigation strategies. Here we show under controlled laboratory conditions that transmission through social networks amplifies motivational biases on a simple artificial decision-making task. Participants in a large behavioural experiment showed increased rates of biased decision-making when part of a social network relative to asocial participants in 40 independently evolving populations. Drawing on ideas from Bayesian statistics, we identify a simple adjustment to content-selection algorithms that is predicted to mitigate bias amplification by generating samples of perspectives from within an individual's network that are more representative of the wider population. In two large experiments, this strategy was effective at reducing bias amplification while maintaining the benefits of information sharing. Simulations show that this algorithm can also be effective in more complex networks.

Large-scale social media platforms are transforming how people communicate and consume information online[1–3], shaping media narratives[4], political discourse[5] and popular culture[6]. However, the complex networks created by social media platforms can have unexpected outcomes. For example, social networks often create 'echo chambers' of like-minded individuals[7–12], raising concerns that social interaction on these platforms can increase polarization and amplify bias[13], facilitating the spread of inaccurate[14–16], extreme[17] and emotionally charged[18] views. Indeed, biased social perceptions can emerge naturally solely from the tendency for network connections to be unevenly distributed[19–21].

Understanding the impacts of social media and developing potential solutions is a global priority, leading to calls for better governance of social media platforms[22,23]. An important part of these potential solutions is the development of content-selection algorithms that demonstrably reduce harmful effects[24]. However, the complexity of large-scale social networks makes it hard to identify specific mechanisms that cause unwanted platform effects and to evaluate the effectiveness of mitigation strategies. These challenges have hindered development of practical frameworks for designing safer algorithms[25].

One way to address these challenges is to test the effects of different content-selection algorithms in controlled laboratory settings. In this Article, we use an experimental paradigm to study how information sharing affects bias in judgement and decision-making. This new experimental paradigm allowed us to evaluate a mathematical theory of bias amplification and test a mitigation strategy based on this theory. Our approach builds on prior research that has demonstrated benefits[26–30] and risks[31,32] of information sharing and that has identified cognitive biases[33] that can be reliably produced in experimental settings. In particular, we examined how social networks can amplify motivated perception, a bias in judgement and decision-making in which people perceive events in a way that supports their desires[34–38].

Our experiments used a simple perceptual task: Participants were briefly shown displays of 100 randomly positioned and sized blue and green dots, and were asked to judge whether the stimulus contained more green or more blue dots (Fig. 1). Participants received a monetary

[1]Department of Psychology, Princeton University, Princeton, NJ, USA. [2]Department of Psychology, University of California, Berkeley, Berkeley, CA, USA. [3]Creative Computing Institute, University of the Arts London, London, UK. [4]Department of Computer Science, Princeton University, Princeton, NJ, USA. [5]These authors contributed equally: Mathew D. Hardy, Bill D. Thompson. ✉e-mail: mdhardy@princeton.edu

https://doi.org/10.1038/s41562-023-01715-5

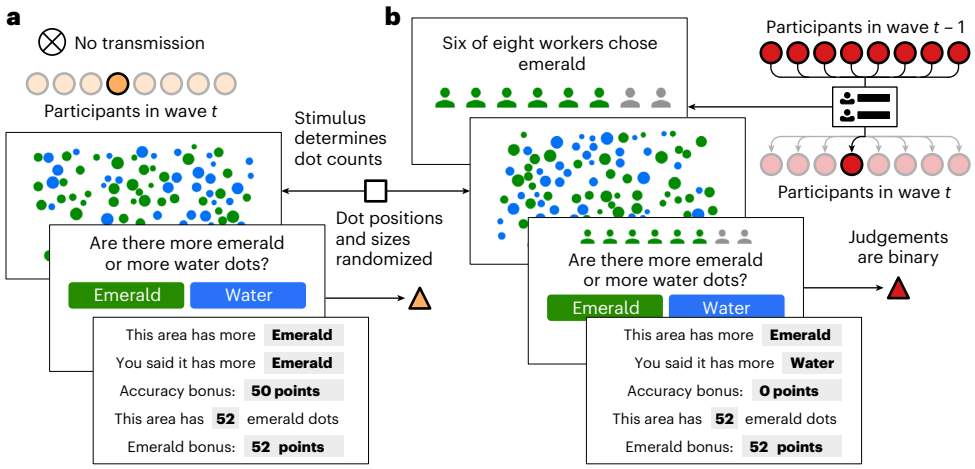


Fig. 1 | Trial structure. Our experimental paradigm consisted of arranging participants into an ordered set of groups, called 'waves'. At each wave *t* participants in social conditions observed judgements made by the participants in wave *t* − 1. Participants in asocial conditions did not observe any social information. **a**, Asocial/Motivated condition. Each coloured circle at the top of the image represents a participant. The square represents the stimulus, and triangles represent the participants' judgements. Stimuli consisted of 100 randomly positioned and sized blue and green dots displayed for one second. After viewing a stimulus, participants indicated whether they thought the stimulus had more green or more blue dots. Participants received feedback after each judgement on practice trials, and at the end of the experiment on test trials. All participants received a bonus on every trial if their judgement was correct. Participants in motivated conditions (shown here) received an additional bonus on every trial for every dot of their motivated colour (green in both plots) regardless of whether their judgement was correct. **b**, Social/Motivated condition. Each coloured circle at the top of the image represents a participant. Solid line arrows indicate flow of information. The perceptual task was the same for participants in asocial and social conditions. However, before and after viewing each stimulus, participants in social conditions observed the aggregate judgements made by participants in the previous wave on the same stimulus. Social information from one wave to the next was presented as the number of participants making the majority judgement (ties were broken randomly).

reward for every correct answer. However, certain participants were offered an additional monetary reward for every green or blue dot in each stimulus ('motivated colour' was randomized across participants; see 'Experiment 1' in Methods). Because it was earned regardless of a participant's judgement, this reward would not influence the judgements of a rational observer. However, on the basis of previous results[35,36] we expected that it would induce a motivated perception bias leading people to overestimate the number of motivated-colour dots. In this context, bias is understood as a tendency to make decisions that are consistent with the rewarded motivation, rather than deviation from the true answer. While artificial, we designed our task so that both of these quantities could be measured, but our approach to bias does not depend on the existence of an objectively true answer.

To better understand how social networks impact decision-making in this controlled setting, we designed custom software to construct evolving social networks by recruiting thousands of participants in sequences of discrete groups, or recruitment waves. This setup builds on a wide body of research in cognitive science and cultural evolution investigating social dynamics in settings with population turnover[31,39–47]. Each wave introduced a new group of participants into the network to replace the previous group of participants. Participants in each wave observed the judgements made by a group of participants in the previous wave, before making their own judgements. This created a layered network of social influence that captured one basic structural feature of online social media: people formed opinions that were informed by metrics summarizing the shared opinions of others. Using this controlled experimental approach, we explored the impact of social networks on biases in decision-making and how unwanted impacts could be mitigated.

To generate quantitative predictions about the consequences of information sharing in this context, we performed a formal analysis of the hypothesis that social networks can amplify people's biases (see 'Bayesian model' in Methods). Our model assumes that social information plays the role of a prior distribution in a Bayesian computation during decision-making, consistent with formal theories of social learning developed in cognitive neuroscience[48], economics[49] and developmental psychology[50]. Our model predicts that participating in a social network can amplify the perceptual signal, potentially increasing accuracy, while also amplifying people's biases (Extended Data Fig. 1). The simultaneous increase in these two quantities is possible because social information effectively amplifies the signal-to-noise ratio of the stimuli as well as the tendency to make bias-aligned judgements. Our model assumes that participants take the observed judgements at face value, without accounting for the fact that the people who made the judgements might also be biased. As a result, bias can compound over time.

Results

Experiment 1 was designed to test the prediction that social networks amplify both accuracy and bias when the ground truth prior probability is uniform and statistically independent from individual people's biases. We then developed a simple mitigation method based on our mathematical analysis, which we evaluated in Experiments 2 and 3.

Experiment 1: bias amplification

The design and results of our first experiment are shown in Fig. 2. Experiment 1 used a factorial design independently manipulating two factors: the presence or absence of the motivated perception intervention ('motivated' conditions versus 'neutral' conditions); and whether participants were shown the judgements of other people or not ('social' conditions versus 'asocial' conditions). Combining these manipulations led to four conditions in Experiment 1: Asocial/Motivated, Asocial/Neutral, Social/Motivated and Social/Neutral. Each new participant in each new wave was assigned to just one of these four conditions using block randomization. Ten independent networks were recruited for each condition, and 64 participants were assigned to each network (8 waves of 8 participants; each participant was assigned to a single network). This design ensured that each condition was entirely independent of the conditions. Altogether, we recruited 2,400 participants from Amazon Mechanical Turk.

https://doi.org/10.1038/s41562-023-01715-5

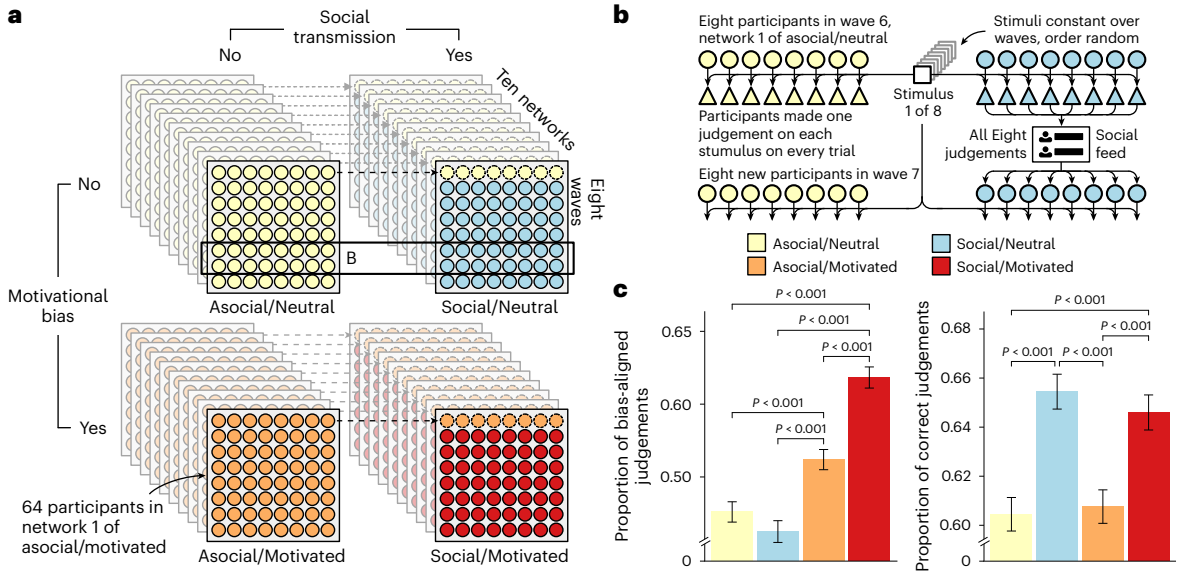


Fig. 2 | Experiment 1 details and results. a, Experiment 1 used a 2 × 2 factorial design, where we varied the presence of social information and of the induced bias in the form of motivated perception. For each condition, we recruited 10 independent networks of 64 (asocial) or 56 (social) participants organized into 8 discrete waves of 8 new participants each. Dotted line arrows indicates a yoking structure we used to reduce variability by controlling the initial settings across the social and asocial conditions (see 'Experiment 1' in Methods; Extended Data Fig. 2). **b**, Experiment 1 transmission structure in the asocial and social conditions. Circles indicate participants, triangles indicate judgements, and squares indicate stimuli. Each stimulus was defined by a number of motivated colour and non-motivated colour dots (see 'Experiment 1' in Methods). The judgements made by social participants at each wave were transmitted to the participants in the same network at the next wave. Asocial participants were recruited in waves to facilitate randomization but did not observe any social information. **c**, Experiment 1 results averaged across waves and networks for each condition (n = 15,120 judgements for asocial conditions; n = 4,480 judgement for social conditions). The plot on the left shows the proportion of participants' judgements that matched their motivated colour. An unbiased participant would choose their motivated colour at a rate of 50% on average. The plot on the right shows the proportion of judgements where participants correctly identified the majority colour of the dots. For both plots, error bars show standard errors of the proportions. P values are determined using a likelihood ratio test between mixed effects models comparing the relevant conditions (P values were not adjusted for multiple comparisons).

To align the starting conditions between asocial and social conditions, social conditions in the second wave were yoked to the relevant asocial conditions in the first wave[47]. That is, Social/Motivated participants in wave 2 viewed the responses of every Asocial/Motivated participant in wave 1 and the same network index, and Social/Neutral participants in wave 2 viewed the responses made by every Asocial/Neutral participant in wave 1 and the same network index (Fig. 2 and Extended Data Fig. 2). In other words, the first wave in the social condition of each network was composed of the first wave from the asocial condition. Participants in social conditions in later waves observed the responses made by participants in the previous wave, condition and network. This initial yoking helped to ensure that any differences in outcome between conditions did not result from differences in the initial judgements.

We designed the task so that motivated perception would not necessarily impact people's ability to perform well: participants made judgements on four stimuli where the correct answer was green, and four where it was blue. Each stimulus was defined in terms of a certain proportion of green and blue dots, but dot positions and sizes were randomized (Fig. 1). Stimulus order was randomized, and participants made one judgement on each stimulus on every trial of the experiment. All participants made judgements on the same eight stimuli (see 'Experiment 1' in Methods). This design led to a sample size of 19,200 judgements: 5,120 for each asocial condition and 4,480 for each social condition. Fewer participants were assigned to the social conditions because all participants in the first wave were assigned to an asocial condition.

To determine whether accuracy and/or bias differed significantly between conditions, we performed a likelihood ratio test between a mixed-effects logistic regression with fixed effects for all conditions and a mixed-effects logistic regression in which the two conditions were coded as a single condition. All regression models included random intercepts for each network. Unless otherwise noted, all analyses and statistical models were pre-registered before collecting human data. Our pre-registered analyses included participant random effects, which we omitted as they led to singular fits. Using the pre-registered models does not change any of our findings (Extended Data Table 1).

As predicted, people in social networks made more accurate judgements. Participants in the Social/Neutral condition identified the correct majority colour on 65.4% of judgements, significantly more often than both Asocial/Neutral (accuracy: 60.4%; β = 0.22, 95% confidence interval (CI) from unrestricted regression: (0.13, 0.3); $\chi^2(1) = 25.65$, $P < 0.001$) and Asocial/Motivated (accuracy: 60.8%; β = 0.2, 95% CI 0.12 to 0.29; $\chi^2(1) = 22.58$, $P < 0.001$). This advantage was also observed in networks in the Social/Motivated condition, as the model predicted. Social/Motivated participants chose correctly on 64.6% of judgements, statistically significantly more often than both Asocial/Neutral participants (β = 0.18, 95% CI 0.09 to 0.26; $\chi^2(1) = 17.6$, $P < 0.001$) and Asocial/Motivated participants (β = 0.16, 95% CI 0.08 to 0.25; $\chi^2(1) = 15.07$, $P < 0.001$). Also as predicted, the colour-based bonus influenced people's judgements. Participants assigned to the Asocial/Motivated group made significantly more biased judgements (skewed towards their motivated colour) than Asocial/Neutral and Social/Neutral participants. Asocial/Motivated participants chose their motivated colour on 56.2% of judgements, compared with 52.6% for Asocial/Neutral participants (β = 0.15, 95% CI 0.07 to 0.22; $\chi^2(1) = 13.36$, $P < 0.001$) and 51.2% for Social/Neutral participants (β = 0.2, 95% CI 0.12 to 0.28); $\chi^2(1) = 23.53$, $P < 0.001$).

Finally, as predicted, motivated perception was amplified among participants who were part of a social network. Social/Motivated

participants chose their motivated colour on 61.8% of judgements, statistically significantly more often than Asocial/Motivated participants ($\beta = 0.23$, 95% CI 0.15 to 0.32; $\chi^2(1) = 31.47$, $P < 0.001$). In Supplementary Information we report exploratory analyses of the mechanisms underpinning these results, showing that information sharing in motivated networks helped people make more accurate judgements specifically when the answer aligned with their biases, and that social influence cannot be reduced to a simpler perceptual priming mechanism (Supplementary Information, 'Exploratory analyses').

We fit our Bayesian model (see 'Bayesian model' in Methods) to data from Experiment 1 using techniques developed in the education literature that exploit the relational structure of testing data[51] (each participant evaluated multiple stimuli, and each stimulus was evaluated by multiple participants; see 'Psychometric model' in Methods). This allowed us to estimate the degree of bias evident in each participant's decision-making while controlling for the ambiguity of each stimulus and the observed social information. As predicted, this model replicates the two main results: that participants' judgements were biased towards their motivated colour; and moderately amplified rates of bias among Social/Motivated participants (Extended Data Fig. 1 and Extended Data Table 2).

Reducing bias amplification

Our analyses illustrate that bias amplification can arise when people observe judgements produced by others who share the same biases. Bias amplification happens because the bias of one wave is passed forward to the next wave through the sample of observed judgements. This suggests that bias amplification could be reduced by adjusting the transmission process so that people observe a set of judgements that better approximate an unbiased sample, representing a broader range of perspectives. Previous research has sought to construct more representative samples by directly curating opposing viewpoints or through heuristic algorithms that alter people's social connections[52–54]. Here we introduce an algorithmic strategy that makes it possible to obtain more representative samples from within an individual's own network. The general form of this problem—approximating a sample from a target distribution p using samples drawn from a different distribution q—can be solved using a method called multiple importance sampling[55,56].

Multiple importance sampling provides a simple formula for assigning a weight w_{jd} to each judgement x_d made by individual j in response to stimulus d. The weight reflects the relative probability of the judgement under the generating distribution $q_j(x_d)$ specific to participant j, and a target sampling distribution $p(x_d)$:

$$w_{jd} \propto \frac{p(x_d)}{q_j(x_d)}. \quad (1)$$

After computing these weights for the judgements made by every individual in each network, a new sample is drawn from this weighted distribution. Intuitively, this algorithm prioritizes judgements that are underrepresented in each individual's network relative to the target distribution.

We used the psychometric model described above to calculate the terms in this weighting procedure (see 'Experiment 2 resampling algorithm' in Methods). Specifically, we calculated q_j using a logistic model with participant-specific bias parameters estimated using hierarchical cognitive modelling at each wave. The target sampling distribution p is the posterior distribution implied by adjusting the model to approximate the entire population.

Following this procedure results in an adjustment to the social information observed by participants on each trial (for example, six of eight people chose green, instead of seven of eight; see 'Experiment 2 resampling scheme' in Methods). In more familiar terms, such an adjustment would modify the algorithm used to select the content presented in a social media feed. The feed would still exclusively contain perspectives drawn from each participant's own network, but the algorithm leverages variation within this network to promote perspectives that are more representative of a broader population. Crucially, this algorithm is information-preserving and does not require knowledge or even existence of a 'ground truth'.

Experiment 2: evaluating resampling

We hypothesized that our importance sampling strategy would mitigate bias amplification while preserving the benefits of information sharing within our experimental paradigm. To test this hypothesis, we evaluated the efficacy of the strategy in a second large experiment. Using the same experimental paradigm as Experiment 1, we recruited 2,464 participants from Mechanical Turk. We determined this sample size by performing a power analysis on 100 simulations of the experiment. These simulations involved alternatively simulating participants' judgements using a model fit to Experiment 1 and simulating the effects of the resampling algorithm on the simulated data (see 'Power analyses' in Supplementary Information).

Participants in Experiment 2 were assigned to one of three conditions: Asocial/Motivated, Social/Motivated and Social/Resampling. The Asocial/Motivated and Social/Motivated conditions replicate Experiment 1. The Social/Resampling condition is analogous to the Social/Motivated condition with the key difference that social information in this condition was subject to the resampling algorithm described above. Participants in social conditions were not informed which social condition they were assigned to. In each condition, half of the networks consisted of participants rewarded for blue dots (blue 'echo chambers'), and half for green dots (green 'echo chambers'). This setup allowed us to impose homophily along motivated colour into each network, which was not present in Experiment 1. We recruited 14 independent networks (7 green and 7 blue) per condition, with 64 participants assigned to each network (8 new participants at each of 8 waves). All participants in the first wave were assigned to the Asocial/Motivated condition. Social/Motivated and Social/Resampling networks were yoked to the same initial asocial judgements. That is, two sets of social participants (Social/Motivated and Social/Resampling) in wave 2 observed the judgements made by the same set of Asocial/Motivated participants in the first wave. At later waves, participants in social conditions observed the responses made by participants in the previous wave and same condition and network. Each participant completed 16 judgements, leading to a sample size of 39,424 judgements: 14,336 for the asocial condition and 12,544 for each social condition. As in Experiment 1, we used an over-recruitment algorithm to speed data collection (more details are given in 'Over-recruitment algorithm' in Methods).

We pre-registered our main analyses and statistical models before collecting data. To test for differences between two conditions, we performed a likelihood ratio test between a mixed-effect logistic regression model with fixed effects for all conditions, and a mixed-effects logistic regression model with the two conditions coded as a single condition. In addition to the condition fixed effects, all models included a fixed effect capturing the participant's motivated colour and random intercepts for each participant.

Replicating Experiment 1, participants in social networks exhibited increases in both bias and accuracy (Fig. 3; full results in Extended Data Table 3). Asocial/Motivated participants chose their motivated colour on 54% of judgements, compared to 56.7% for Social/Motivated participants ($\beta = -0.12$, 95% CI −0.19 to 0.05); $\chi^2(1) = 12.53$, $P < 0.001$). Similarly, Asocial/Motivated participants chose correctly on 57.4% of trials, significantly less often than Social/Motivated participants, who chose correctly on 63% of judgements ($\beta = -0.23$, 95% CI −0.29 to −0.18); $\chi^2(1) = 82.54$, $P < 0.001$). As predicted, however, the resampling adjustment mitigated bias amplification. Social/Resampling participants chose their motivated colour on 54.5% of judgements, a significantly lower rate than Social/Motivated participants ($\beta = -0.1$, 95% CI −0.16

https://doi.org/10.1038/s41562-023-01715-5

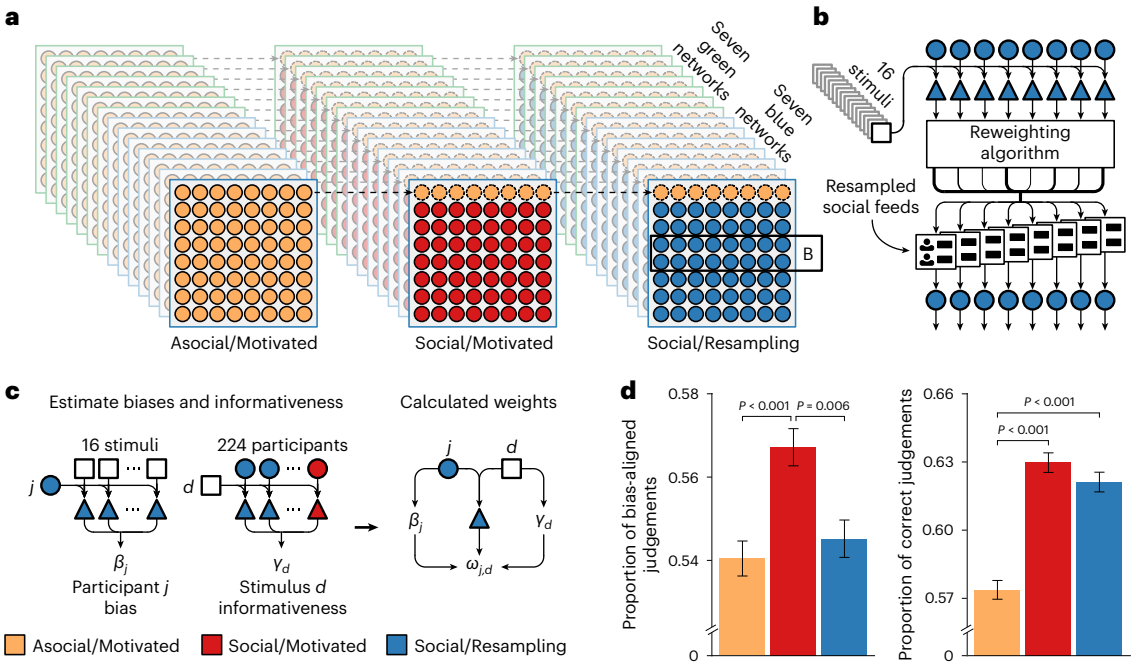


Fig. 3 | Experiment 2 details and results. a, Experiment 2 had three conditions: Asocial/Motivated, Social/Motivated and Social/Resampling. For each condition, we recruited 14 independent networks of participants: 7 networks of participants paid for green dots, and 7 networks of participants paid for blue dots. As in Experiment 1, participants in each network were organized into eight discrete waves, with eight participants in each wave. Dashed lines illustrate the yoking procedure, where second-wave participants in both social conditions observed the judgements made by first-wave Asocial/Motivated participants. **b**, The transmission structure for a single network in the Social/Resampling condition. At the end of each wave, we assigned a weight to each judgement made by Social/Resampling participants in that wave. This weight determined the probability of propagating the judgement to participants at the next wave. An independent resampling of judgements was done using the same weights for each participant in the next wave. **c**, Calculating the resampling weights for each judgement was a two-step process. We first used a psychometric model to jointly estimate the bias (β) of each participant j and the 'informativeness' γ_d of each stimulus d. This model was fit to Asocial/Motivated participants in wave 1, and to Social/Motivated and Social/Resampling participants in waves 2–8. Biases were estimated from all of a participant's 16 judgements. Informativeness for a stimulus d was estimated using all 224 judgements the model was fit to on that stimulus. We then determined the weight for propagating each judgement using equation (1). **d**, Experiment 2 results averaged across waves and networks (n = 14,336 judgements for the asocial condition; n = 12,544 judgements for each social condition). The plot on the left shows the proportion of trials where participants' judgements corresponded to their motivated colour, and the plot on the right the proportion of trials where participants chose correctly. Proportions are averaged across all participants in each condition and error bars show standard errors of the proportions. P values are determined using a likelihood ratio test between mixed effects models comparing the relevant conditions (P values were not adjusted for multiple comparisons).

to −0.03); $\chi^2(1) = 7.47$, $P = 0.006$). Furthermore, there was no evidence that resampling reduced the accuracy benefit of social networks. Social/Resampling participants chose correctly on 62.1% of judgements, significantly more often than Asocial/Motivated participants ($\beta = 0.2$, 95% CI 0.15 to 0.25); $\chi^2(1) = 59.33$, $P < 0.001$), and not statistically significantly different from Social/Motivated networks ($\beta = -0.04$, 95% CI −0.09 to 0.02; $\chi^2(1) = 1.87$, $P = 0.172$). In Supplementary Information, we report exploratory analyses showing that our algorithm reduced the bias of observed social information in both blue- and green-biased networks, facilitating greater consensus (Extended Data Fig. 3; 'Experiment 2 exploratory analyses' in Supplementary Information). The algorithm did so while propagating judgements from each participant at similar rates, rather than for example only transmitting judgements made by participants with low estimated bias (Extended Data Fig. 3).

Experiment 3: evaluating resampling in skewed networks

The results of Experiment 2 suggest that resampling can be effective in settings where individuals observe representative samples of judgements from their own networks. Is resampling also effective in settings in skewed networks where more biased individuals are more likely to share their judgements? This phenomenon has been observed to occur in social media[57], and potentially increases the bias of the population.

To investigate, we performed a third experiment using a similar setup as Experiments 1 and 2. Specifically, we tested the effectiveness of our algorithm in networks where the probability that a participant j has their judgement propagated to others at the next wave is proportional to the magnitude of their bias h_j (for more details, see 'Experiment 3' in Methods).

Participants were recruited in two discrete waves and were assigned to one of three conditions; Asocial/Motivated, Social/Correlated and Social/Resampling. As in Experiments 1 and 2, both waves consisted of different participants. All participants in the first wave were assigned to the Asocial/Motivated condition, all participants in the second wave to either the Social/Correlated or the Social/Resampling condition. The Asocial/Motivated replicates the earlier experiments. In the Social/Correlated condition, participants observed a set of judgements from participants in the previous wave. This set of judgements was sampled with replacement from the set of eight judgements that would be propagated in the Social/Motivated condition, with the probability of sampling each judgement proportional to the magnitude of an estimate of that participant's bias (see 'Experiment 3' in Methods). That is, participants with high levels of bias towards either green or blue would tend to have their judgements transmitted at higher rates. In Social/Resampling networks, we used our resampling procedure to adjust this skewed set of judgements with the goal of obtaining a more representative sample.

We recruited 320 participants for each condition from Mechanical Turk, for a total of 960 participants. We determined this sample size by performing a power analysis on 100 simulations of the experiment. As in Experiment 2, these simulations involved alternatively simulating

https://doi.org/10.1038/s41562-023-01715-5

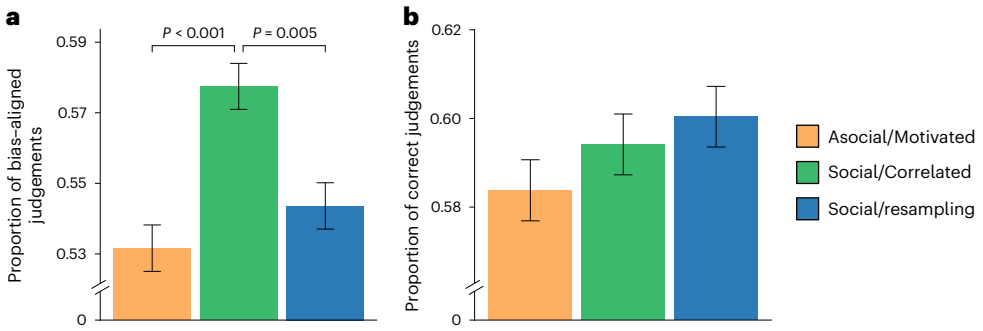


Fig. 4 | Experiment 3 results. a, Plot of the proportion of trials where participants' judgements corresponded to their motivated colour. **b**, Plot of the the proportion of trials where participants chose correctly. Participants in the Social/Correlated condition observed a set of judgements from participants in the previous wave. These judgements were sampled on the basis of the magnitude of each participant's bias, leading to a higher transmission rate of judgements from more biased individuals. Participants in the Social/Resampling condition observed a set of judgements resampled from the original skewed set using our resampling algorithm. As predicted, we found that resampling significantly reduced bias amplification. Proportions are averaged across all participants in each condition (n = 5,120 judgements per condition). Error bars show standard errors of the proportions. P values are determined using a likelihood ratio test between mixed effects models comparing the relevant conditions (P values were not adjusted for multiple comparisons).

participants' judgements using a model fit to Experiment 1 and simulating the effects of the oversampling and resampling algorithm on the simulated data (see 'Power analyses' in Supplementary Information).

All participants were assigned to one of 40 independent networks. Half of the networks consisted of participants paid for blue dots, and half for green dots, with eight participants in a network at each wave. Each participant made judgements on the same 16 test stimuli, matching the proportions used in Experiment 2.

Both social conditions were yoked to the same initial asocial data. For example, participants in the network with index n in both the Social/Correlated condition and Social/Resampling conditions observed judgements sampled from the same set of judgements made by Asocial/Motivated participants' in the network with index n.

Our sample size, analyses, and statistical models were pre-registered before collecting data. Our pre-registered analyses included random effects for each social network to account for the yoking structure, which we omitted as they led to singular fits. Using the pre-registered models does not change any of our findings (for full results, see Extended Data Table 4).

As in Experiments 1 and 2, we compared conditions by performing a likelihood ratio test between two logistic regression models: one model with both conditions coded separately, and one where they were coded as the same condition. All regression models also included fixed effects for motivated colour and random intercepts for each participant. Tests that compare bias used models that predicted whether participants chose their motivated colour, and tests that compare accuracy predicted a binary variable capturing whether participants chose correctly. Results are shown in Fig. 4.

As predicted, and mirroring the results from Experiments 1 and 2, we found that participants in the Social/Correlated condition made significantly more biased judgements that participants in the Asocial/Motivated condition—Social/Correlated participants chose their motivated colour on 57.6% of trials, compared with 52.7% of trials of Asocial/Motivated participants (β = 0.21, 95% CI 0.1 to 0.32; $\chi^2(1) = 14.7$, $P < 0.001$). However, our resampling algorithm significantly reduced bias amplification; Social/Resampling participants chose their bias colour on 54.0% of trials, significantly less than Social/Correlated participants (β = −0.15, 95% CI −0.26 to −0.05); $\chi^2(1) = 8.0$, $P = 0.005$). Furthermore, as in Experiment 2, we found no statistically significant reduction in accuracy between Social/Correlated and Social/Resampling participants; Social/Correlated participants chose correctly on 59.4% of trials, and Social/Resampling participants chose correctly on 59.9% of trials (β = 0.03, 95% CI −0.06 to 0.11; $\chi^2(1) = 0.4$, $P = 0.535$).

Network simulations

Our main analyses and experiments investigated networks in which information is passed through discrete waves of participants organised in completely connected layered networks. However, previous research has demonstrated important impacts of network structure on how information, beliefs and behaviours spread through populations[58–63]. To assess the robustness of our method, we used the model we fit to participant behaviour in Experiment 1 to conduct simulations measuring the effects of resampling in more complex networks with repeated interaction.

Similar to the power analyses conducted for Experiments 2 and 3, we simulated behaviour in networks by alternatively simulating participants' judgements using a model fit to experimental data, and then simulating the impact of our resampling algorithm using models fit to the simulated data (Extended Data Fig. 4). We generated random networks using a power-law distribution to determine the number of outgoing edges for each node. This distribution was chosen to model inequality in social networks, where a few nodes (that is, influencers) tend to have a disproportionately high number of connections. The power law distribution is defined by a parameter a that determines the degree of inequality in the network. Lower values of a lead to more unequal networks—as a decreases, a few nodes, or 'hubs', tend to amass a disproportionately large number of outgoing edges, resulting in a skewed network structure.

We compared networks with resampling to those without for three different network sizes (64, 128 and 256) and three levels of a (1.5, 2.0 and 3.0). We simulated 100 network sets (green and blue Social/Motivated and Social/Resampling networks) for each unique combination of parameters (for more details, see 'Network simulation' in Methods).

In line with our main results, we found that resampling significantly reduced bias amplification for every a and network size we tested (for results, see Extended Data Table 5 and Extended Data Fig. 5). Furthermore, resampling significantly increased accuracy for every network type, although the magnitude of the effect was small.

Discussion

Together, these results help to identify the mechanisms contributing to bias amplification in social networks. This study recreated a simplified version of bias amplification in a controlled laboratory experiment where we could directly measure the effects of information sharing on bias by comparing the judgements made by isolated and connected individuals. Experiment 1 showed that information sharing led to the amplification of a decision-making bias beyond the levels of bias expressed by individuals completing the task alone. Experiment 2

showed that this amplification effect can be mitigated by a statistical resampling strategy applied to the transmission process. Experiment 3 showed that resampling can be effective in settings where bias is correlated with sharing, and simulations showed that the method can also be successful in more complex networks.

The methods we have presented here—using psychometric models to estimate bias and provide a quantitative foundation for resampling—can be extended to more naturalistic settings where the relative probability of each response under a personalized generating distribution and a target sampling distribution can be evaluated. For example, if links that are shared by users of a social media platform can be identified as having a liberal or conservative bias, this information can be used to estimate the biases of individuals who choose to share those links. Resampling approaches may also be effective in reducing biased perceptions that arise from 'friendship-paradox' phenomena where people's perceptions are overly influenced by highly connected individuals[64].

While promising, further work is needed to better understand the robustness of our resampling scheme and modelling framework. Most notably, the networks we constructed for our experiments were both simple and relatively small. Real-world social networks, by contrast, are both larger and more complex[65–67]. While simulations suggest resampling can be effective in more complex networks, future work should test this experimentally, especially in settings where sharing is mediated by algorithms that prioritize engagement. Along these lines, in real-world networks, connections arise from a variety of complex structures and relationships. Future work should investigate whether resampling can also be effective in networks where connections are formed through these richer processes. Adjusting social information may also be less successful at influencing people's beliefs on complex topics that interact with people's identities and experiences.

Our modelling framework also made a number of simplifying assumptions that may need to be adjusted in real-world domains. For example, for simplicity we assume that individuals do not account for the bias of the social information they observe or whether others are also social learners. This assumption may break down in naturalistic settings where a group's identity, history, goals and biases are more explicit, and previous work has shown that people can be sensitive to similar dependencies[68]. Investigating the extent to which individuals account for the generative process of social information both in laboratory tasks and real-world networks is an exciting challenge for future work, alongside further investigation of principles for maintenance of sample diversity in more complex decision-making settings, and integration of our approach with other theoretical models of social learning[49,69].

Online social networks are often constructed by opaque algorithms that prioritize user engagement[70,71]. We have identified and evaluated a simple adjustment to social networking algorithms that reduced bias amplification while maintaining the benefits of information sharing. Resampling allowed participants to observe social information that was generated by people in their network but that better reflected the broader population. This method complements interventions such as fact-checking[72], education initiatives[73] and content moderation[74], providing a framework for reducing bias amplification that is transparent, scalable and underpinned by a mathematical theory that shows why the approach is effective.

Methods

Bayesian model

Our Bayesian model assumes that participants face a decision between two hypotheses: more green dots ($h = g$) or more blue dots ($h = b$). The evidence on which this judgement is based is the observed data d and, optionally, social information s in which k of n other participants endorse green. We thus want to compute $p(h = g|d, s)$. With just two hypotheses, we can write Bayes' rule in log-odds form:

$$\log \frac{p(h=g|d,s)}{p(h=b|d,s)} = \log \frac{p(s|h=g)}{p(s|h=b)} + \log \frac{p(d|h=g)}{p(d|h=b)} + \log \frac{p(h=g)}{p(h=b)} \quad (2)$$

where we assume that the data d and social information s are independent, with $p(d, s|h) = p(d|h)p(s|h)$, and $p(h)$ reflects prior biases.

For the social information s, we assume that participants assume that each of the judgements is generated independently with probability $1 - \epsilon$ of matching the true colour (note that the true generative process depends on both the wave and others' motivated colour). Under these assumptions, social information in which k of n judgements favour green thus results in

$$\frac{p(s|h=g)}{p(s|h=b)} = \frac{(1-\epsilon)^k \epsilon^{n-k}}{\epsilon^k (1-\epsilon)^{n-k}} \quad (3)$$

$$= \left(\frac{\epsilon}{1-\epsilon}\right)^n \frac{(1-\epsilon)^k \epsilon^{-k}}{\epsilon^k (1-\epsilon)^{-k}} \quad (4)$$

$$= \left(\frac{\epsilon}{1-\epsilon}\right)^n \left(\frac{1-\epsilon}{\epsilon}\right)^{2k} \quad (5)$$

$$= \left(\frac{1-\epsilon}{\epsilon}\right)^{2(k-n/2)} \quad (6)$$

which allows us to write

$$\log \frac{p(s|h=g)}{p(s|h=b)} = (k - n/2) \log \left(\frac{1-\epsilon}{\epsilon}\right)^2.$$

Using this result, we can rewrite equation (2) as

$$\log \frac{p(h=g|d,s)}{p(h=b|d,s)} = \alpha(k - n/2) + \gamma_d + \beta,$$

where

$$\alpha = \log \left(\frac{1-\epsilon}{\epsilon}\right)^2 \quad (7)$$

$$\gamma_d = \log \frac{p(d|h=g)}{p(d|h=b)} \quad (8)$$

$$\beta = \log \frac{p(h=g)}{p(h=b)}. \quad (9)$$

By observing that

$$p = \frac{1}{1 + \exp\left\{-\log \frac{p}{1-p}\right\}},$$

and substituting $p(h = g|d, s)$ for p, we can derive a logistic expression for judging green[75]:

$$p(h=g|d,s) = \frac{1}{1 + \exp\{-\alpha(k - n/2) - \gamma_d - \beta\}}. \quad (10)$$

To model social network dynamics, we define a Markov chain on the number of people endorsing green at time t, k_t, for a specific stimulus d. The transition matrix of the Markov chain is specified by

$$k_t \sim \text{Binomial}\left(n, \frac{1}{1 + \exp\{-\alpha(k_{t-1} - n/2) - \gamma_d - \beta\}}\right).$$

Since γ_d and β behave equivalently in this model, it is sufficient to examine how the stationary distribution of the Markov chain is affected by β. Extended Data Fig. 1 shows the stationary distribution as a function of β for three different values of α, reflecting different levels of impact of social information. At all levels of α, the stationary distribution exaggerates the effect of β (and hence γ_d), with this phenomenon increasing as α increases. Intuitively, this is a form of compounding: β and γ_d have an effect on judgements at each wave, and the social information conveys that effect to the next wave where it has a further opportunity to influence people's judgements. In the case of γ_d this potentially increases accuracy, as people make better use of limited information—a positive effect of social networks. However, for β this is simply a magnification of bias.

Experiment 1

Participants. All experiments were approved by the institutional review board of Princeton University, and all participants gave informed consent. We recruited 2,619 participants from Amazon's Mechanical Turk, limiting our study to those located in the United States. We only recruited participants with a Mechanical Turk approval rating of 95% or higher, and participants could only take the experiment once. Individuals that participated in a pilot version of the study were excluded from the experiment. Participants received a base payment of US$0.65 for completing the experiment, plus an average bonus of US$0.65. On average, participants spent 4 min and 10 s on the task, and earned an average hourly wage of US$21.73. Our recruitment algorithm used planned over-recruitment to facilitate efficient data collection in the context of sequential batch recruitment, enabling us to recruit the target sample size of 2,400 participants (see 'Over-recruitment algorithm' in Methods). Over-recruited participants were paid using the same procedure, but were excluded from analyses and did not have their judgements transmitted to the next wave. Our sample size and analyses were pre-registered on 19 September 2019 (pre-registration available at ref. 76).

Conditions. Our experiment used a 2 × 2 factorial design with two binary factors; presence of motivation bias (induced by the experimenters) and availability of social information. Each participant was randomly assigned to one of the four conditions implied by this cross.

The reward structure factor had two levels: motivated and neutral. In motivated conditions, participants were randomly assigned a motivated colour of either green or blue. These participants received one point for every dot of their motivated colour on each trial. To ensure the expected total reward for participants in all conditions was equal, participants in neutral conditions received an additional 400 point reward for completing the experiment. Participants in all conditions also received a 50 point reward on each trial if their judgement was correct. At the end of the experiment, participants' bonuses were paid to them with 10 points equal to one cent.

The social factor also had two levels: asocial and social. Participants in the social conditions observed a set of other participants' responses on the same stimulus before each trial. Participants in asocial conditions performed each trial individually.

Wave structure. Participants were recruited in discrete batches, or 'waves'. A wave of participants was not recruited until all the participants in the previous wave completed the experiment. All participants in the first wave were randomly assigned to one of the two asocial conditions, and participants in later waves were randomly assigned to one of the four conditions.

Network structure. We recruited ten independent networks for each condition. After being assigned to a condition, each participant was then randomly assigned to a network. Each network had eight participants at each wave. Network determined the responses participants in a given wave observed. That is, the network captured the social network the participant was assigned to. This network was constant across trials, so a participant observed the responses by the same set of participants on each trial. Participants in asocial conditions were assigned to a network but did not observe any social information.

We created networks in which participants were rewarded for different colours. To do so, each participant was assigned a marked colour of either green or blue. Equal numbers of participants in each network and condition were assigned each colour. In the motivated conditions, marked colour corresponded to the participant's motivated colour. Participants in neutral conditions, however, were not aware of their marked colour. Response data were coded as whether the participant chose their marked colour, and the judgements observed by a participant were recoded in terms of that participant's marked colour (Extended Data Fig. 2).

Procedure. All practice and test stimuli consisted of 100 randomly positioned blue and green dots displayed for 1,000 ms. After viewing a stimulus, participants judged whether the image contained more blue or more green dots. The blue (#007ef8, LAB: 53.49, 16.3, −69.16) and green (#009500, LAB: 53.49, −57.73, 55.72) colours were chosen for identical lightness in LAB colour space. A fixation cross and bounding box preceded presentation of the stimuli for 600 ms.

For each stimulus, the position and size of dots varied between networks and waves, but not across conditions. That is, for each network, participants in condition A observed with the exact stimuli viewed by participants in condition B. For example, all participants in wave t and network r observed the exact same display ($S_{t,r}$) for each stimulus. Participants in the next wave but same network making judgements on the same stimulus observed a display ($S_{t+1,r}$) with the same number of marked colour dots as $S_{t,r}$, but that were configured in different sizes and positions. A participant's marked colour then determined the colour of each dot.

Each participant first completed two practice trials. Participants received accuracy and reward feedback after each practice trial. Participants were then required to pass a comprehension test in their first three tries to complete the experiment. Each participant completed eight test trials. No trial-by-trial feedback was given on test trials. Instead, participants were told their earnings and performance at the end of the experiment.

Each practice and test stimulus was associated with a number of marked and non-marked colour dots. Practice stimuli had 47 and 53 marked-colour dots. Test stimuli had 48, 49, 51 or 52 marked-colour dots, with two trials for each unique number. Trial order for both practice and test trials was randomized.

Social information was presented as the number of participant making the majority judgement. For example, if six of eight participants in the previous wave chose their marked colour, then a participant P at the next wave whose marked colour was blue was informed that six of eight participants in the previous wave chose blue. However, if only three of eight participants chose their marked colour, then P was informed that five of eight participants chose green. When equal numbers of participants chose their marked and non-marked colours, we used simple randomization to determine which colour was presented as the response of four out of eight individuals. The social icons observed before the stimulus were presented again above the blue and green response buttons when participants made their judgement.

We used cover stories to make the task more intuitive. All participants were informed that they were working for an imaginary mining company looking for valuable gemstones. Participants in motivated conditions whose marked colour was blue were told that they would be looking for blue sapphires in green grass, and to inspect areas and judge if the area had more sapphire dots or more grass dots. Participants in motivated conditions whose marked colour was blue were given a similar cover story, but were instructed that they were looking

for green emeralds in blue water. Participants in neutral conditions were told to judge whether the area contained more blue sapphires or green emeralds.

Statistical testing. Differences in accuracy and bias between conditions were assessed using likelihood ratio tests between relevant mixed effects models (Extended Data Table 1). The likelihood ratio test assumes a sample size large enough for the test statistic to follow a chi-square distribution and independent observations. We did not formally test these assumptions for our Experiment 1 analyses or any subsequent likelihood ratio test reported in this paper.

Psychometric model

We fitted separate versions of the Bayesian model introduced above to the judgements made by Asocial/Motivated and Social/Motivated participants in Experiment 1. These models can also be interpreted in terms of Item Response Theory[51], a psychometric approach typically used to jointly estimate individuals' abilities and questions' difficulties using testing data. We used this approach to estimate the impact of social information α, the informativeness of the data γ_d and the biases of participants β.

The model specified in equation (10) in the main paper is a form of logistic regression. We fit two models—one for the Social/Motivated condition and one for the Asocial/Motivated condition—via Bayesian logistic regression. We estimated the bias b_j of participant j towards their motivated colour using a hierarchical Bayesian model, with $b_j \sim \mathcal{N}(\mu_b, \sigma_b)$, $\mu_b \sim \mathcal{N}(0, 3)$ and σ_b ~ lognormal(0, 2) (fitting the model with different priors results in minimal changes to the fitted parameters). This model allows us to pool data across participant groups and increases statistical power. The estimate μ_b of the mean of the population-level distribution captures the tendency of blue- and green-biased participants to choose their motivated colour. This can be then recoded as the bias β_j of participant j towards green in the regression model:

$$\beta_j = \begin{cases} b_j, & \text{if } j\text{'s motivated colour is green} \\ -b_j, & \text{if } j\text{'s motivated colour is blue.} \end{cases} \quad (11)$$

To model each stimulus, both models included coefficients $\boldsymbol{\gamma_d} \sim \mathcal{N}(0, 20)$ on each level of a dummy variable representing the number of green dots (0.48, 0.49, 0.51 and 0.52) in the stimulus. We modelled the impact of the social information s on Social/Motivated participants' judgements by fitting a weight $\alpha \sim \mathcal{N}(0, 20)$ on $k - n/2$, or the number of green judgements observed for the stimulus minus 4 ($n/2$). The asocial model did not include α. Controlling for this social information allowed us to compare the individual-level bias of Asocial/Motivated and Social/Motivated participants using the bias term μ_b as described above.

Parameters in both models were estimated from eight Markov chain Monte Carlo chains run for 2,500 iterations using the No-U-Turn variant of Hamiltonian Monte Carlo[77]. The first 1,250 iterations of each chain were discarded and not used for parameter estimation. Average sampled values and highest density intervals for both models are given in Extended Data Table 2.

Experiment 1 model fitting. We used this model to analyse the judgements made by Asocial/Motivated and Social/Motivated participants. In both models, the mean bias towards motivated colour (β for participants paid for green dots, $-\beta$ for participants paid for blue dots) was positive and significant, indicating that participants' judgements were skewed towards their motivated colour (Asocial/Motivated bias mean 0.268, 90% highest density interval 0.218–0.32; Social/Motivated bias mean 0.371, 90% highest density interval 0.302–0.44; full results in Extended Data Table 2). Similarly, there was a monotonic relationship between the proportion of green dots and the estimated coefficient γ_d for each stimulus. Finally, in the Social/Motivated model, we observed a positive relationship between the number of prior judgements of green and the probability of choosing green ($\alpha = 0.232$, 90% highest density interval 0.199–0.265). These parameters are predicted to lead to moderately amplified bias (Extended Data Fig. 1).

Experiment 2 resampling algorithm

The resampling algorithm used in Experiment 2 involved (1) fitting a psychometric model to estimate participant biases and stimulus informativeness, (2) estimating weights for each judgement using multiple importance sampling, and (3) resampling a set of judgements to propagate to participants at the next wave. Here we describe each step of this process in detail.

Estimating β and γ_d. We modelled the generating distribution q and the unbiased target distribution p using a psychometric model based on the Bayesian model introduced above. Participant biases were assumed to be $\beta_j \sim N(0, 3)$ and stimulus informativeness was $\gamma_d \sim N(0, 3)$. We omitted α from this version of the model. We did this to establish that the resampling algorithm can be applied in the most general case where the observation history of the people whose judgements are being resampled is unknown. To estimate q (the probability of the judgement under the generating distribution) we use a Bayesian logistic regression model that incorporates both of these factors. To estimate p (the probability of the judgement under the target distribution), we calculate the distribution that results from setting $\beta_j = 0$. This approach is an appropriate model of the unbiased broader population in our setting because there were equal numbers of participants with green ($\beta > 0$) and blue ($\beta < 0$) biases, allowing us to assume that the mean of this distribution is zero; in other settings, the target β can be set to the empirical population mean. Unlike the Experiment 1 psychometric model, coefficients were estimated for each stimulus directly from judgement data. That is, we fit γ_d without using the true dot proportions (that is, the correct answer). Additionally, our approach does not require knowledge of the true bias of any individual. Instead, our framework allows us to estimate biases through distributional assumptions about the population.

At the end of each wave, we fit the item response theory (IRT) model by running eight Markov chain Monte Carlo chains for 2,500 iterations (with the first half warmup) using the No-U-Turn variant of Hamiltonian Monte Carlo[77]. In the first wave, the model was fit to Asocial/Motivated participants, and in later waves it was fit to Social/Motivated and Social/Resampling participants.

Multiple importance sampling. We aimed to obtain an unbiased sample of judgements while only having access to judgements produced by a biased population. This problem—approximating a sample from a target distribution p using samples drawn from a different distribution q—can be solved with importance sampling[55]. In importance sampling, a weight $\tilde{w}_i$ is assigned to each sample x_i, with:

$$\tilde{w}_i = \frac{p(x_i)}{q(x_i)}. \quad (12)$$

This setup assumes that all observations are drawn from a common generating distribution q. In our setting, however, bias can vary across individuals. This means that the observed judgements are drawn from a distinct generating distribution q_j for each participant j. That is, we estimate the probability of the judgement on stimulus d using both the stimulus parameter γ_d and the fitted bias parameter β_j for participant j. For example, all else equal, q_j will be lower on judgements where participants with a high bias towards green chose blue.

This setting requires a more specific form of adjustment known as multiple importance sampling[56]. One method for performing multiple

importance sampling is to set the weight of each observation equal to its probability under p divided by its probability under q_j:

$$\tilde{w}_i = \frac{p(x_i)}{q_j(x_i)}. \quad (13)$$

The weights are then normalized so they can be interpreted as components of a probability vector. This can be done by dividing each weight by the sum of the weights for all N observations:

$$w_i = \frac{\tilde{w}_i}{\sum_{n=1}^{N} \tilde{w}_n}. \quad (14)$$

Resampling judgements. In our setting, q is the distribution of biased judgements and p is the unbiased distribution. Given a sample from q, we can use importance sampling to draw a sample of judgements that is closer to p. As described above, this procedure involves assigning weights to the original set of judgements and then sampling a new set of judgements from the set with probabilities proportional to the weights.

At the end of each wave, we used the estimated IRT models to calculate p, the probability of the judgement under an average bias towards green, and q, the probability of the judgement given the participant's bias for all judgements in the wave. We used equation (13) to calculate the unnormalized weight $\tilde{w}_{jd}$ for every judgement made by Asocial/Motivated (wave 1) or Social/Resampling (waves 2–8) participants. To normalize the weights, we assigned each judgement to a judgement set. A judgement set consisted of the $N = 8$ judgements made by the participants in the same network and wave on stimulus d. Following equation (14), each weight was normalized by dividing it by the sum of all the weights in the judgement set.

The normalized weights $\boldsymbol{w}$ determined the probabilities of propagating each judgement to participants at the next wave. On every Social/Resampling trial, we used these probabilities to sample a set of eight judgements from the relevant judgement set. Judgements were resampled with replacement, and distinct samples were propagated to each participant at the next wave.

Experiment 2

Participants. All experiments were approved by the institutional review board of Princeton University, and all participants gave informed consent. We recruited 2,518 participants from Amazon's Mechanical Turk. As in Experiment 1, we required participants to be based in the United States and have an approval rating of at least 95%. To keep up to date with evolving data quality norms on Mechanical Turk[78] (Experiment 1 was run in September 2019; Experiment 2 was run in November 2021), we also required participants to have successfully completed at least 5,000 human intelligence tasks. Participants could only take the experiment once. We allowed individuals who completed Experiment 1 or its pilot studies to participate in the experiment. However, individuals that participated in a pilot version of Experiment 2 were excluded. Participants received a base payment of US$1.20 plus an average bonus of US$1.28. Participants spent an average of 5 min and 24 s on the task, and earned an average hourly wage of US$31.82. As in Experiment 1, we used an over-recruitment algorithm to accelerate collection of our target sample size of 2,464 participants. Over-recruited participants were paid using the same procedure, but were excluded from analyses and did have their judgements transmitted to the next wave. Our sample size and analyses were pre-registered on 2 August 2021 (pre-registration available at ref. 79).

Conditions. All participants were assigned a motivated colour for which they received a one point bonus for every dot of that colour. In addition to their colour bonus, all participants received a 50 point bonus on each trial if their judgement was correct. At the end of the experiment, participants' bonuses were paid to them with 10 points equal to one cent.

Participants were assigned to one of three conditions—Asocial/Motivated, Social/Motivated and Social/Resampling. Participants in the asocial condition completed the experiment individually and did not observe social information. Social participants observed a set of judgements made by participants in the previous wave on the same stimulus before making a judgement. Social/Motivated participants observed the original judgement set. Participants in Social/Resampling networks observed a set of eight judgements sampled with replacement from the judgement set. We resampled a separate set of eight judgements for each participant.

Wave structure. As in Experiment 1, participants were recruited in eight discrete waves. Networks and conditions were run in parallel, so participants at wave $t + 1$ were not recruited until all participants in wave t completed the experiment. All participants in the first wave were assigned to the asocial condition. Equal numbers of participants in later waves were assigned to each condition using block randomization.

Network structure. At each wave, participants in each condition were randomly assigned to a network. We recruited 14 networks for each condition. A network determined the responses social participants observed. This network was fixed, so a social participant observed social information from the same set of participants on each trial.

As in Experiment 1, each network had eight participants in each wave. Instead of assigning participants with both green and blue biases to the same network as in Experiment 1, networks were composed entirely of participants with the same motivated colour. Seven of the networks for each condition consisted of participants biased towards green, with the other seven populated by participants biased towards green. This structure allowed us to avoid marked-colour recoding within networks as we did in Experiment 1.

Procedure. As in Experiment 1, all stimuli consisted of 100 dots displayed for one second. Participants then judged whether there were more blue or more green dots in the stimulus. The blue (#007ef8) and green (#009500) colours matched those used in Experiment 1. A fixation cross and bounding box was displayed for 600 ms before presentation of the stimuli. Positioning and sizing for a given stimulus was determined randomly under the constraint that no dots overlap with each other or the stimulus boundary. Unlike Experiment 1, we did not constrain dot positioning and sizing to match on each wave and network.

Social information on each trial was shown before participants in social conditions viewed a stimulus. Social information consisted of both text and icons, with k icons coloured to match the judgement of the majority, and $8 - k$ coloured in grey. For example, six green and two blue judgements were displayed as six green and two grey icons. However, if only three of eight participants chose green, participants were shown five blue icons and three grey icons. Ties were broken via simple randomization. The same social icons a participant observed before a stimulus were also presented above the blue and green response buttons when participants made their judgement.

Participants first completed two practice trials with 53 and 47 green dots. Participants received accuracy and reward feedback after both practice trials. Practice trials did not count towards participants' bonuses, and no resampling was performed on practice trials in Social/Resampling networks. Participants then completed a comprehension test before starting test rounds. Participants that failed to answer every test question correctly in their first three tries were excluded from the experiment.

As in Experiment 1, test stimuli had either 48, 49, 51 or 52 green dots (our pre-registration erroneously stated that test stimuli would have either 49 or 51 green dots). Participants completed 16 test trials, with

4 test trials for every number of green dots. No trial-by-trial feedback was given on the test trials. Instead, participants were informed of their earnings at the end of the experiment. Trial order for both practice and test trials was randomized.

To make the task more intuitive, we used the same cover stories we used in Experiment 2. Participants were told that they were working for an imaginary mining company looking for valuable gemstones. Participants whose motivated colour was blue were instructed that they were looking for blue sapphires in green grass, and those whose motivated colour was green were told that they were looking for green emeralds in blue water.

Over-recruitment algorithm

In Experiments 1 and 2, we applied an over-recruitment procedure designed to facilitate efficient network turnover in sequential, batch-structured recruitment settings where a full wave must complete before the next wave can be started (we did not use over-recruitment in Experiment 3). Specifically, at each wave, we performed planned over-recruitment at the experiment level. That is, we first calculated the required total sample size for the current wave including all networks. We then added a number of additional potentially fillable recruitment slots to that total. This ensured that participants who began the task but failed to complete it or took a disproportionately long time to do so did not prevent completion of a wave, which occurred frequently during pilot testing.

As soon as the target sample size for a wave had been completed, the new wave was recruited. Any participants who were in the process of completing their task when their wave hit the target sample size completion were automatically assigned to an over-recruited category. Participants assigned to the over-recruited category completed the task normally and received the same payments and bonuses as all other participants, but their data were not transmitted to the next wave nor included during our analyses.

Our procedure aimed to keep the number of over-recruitment slots available during the experiment constant. For example, in Experiment 1, our total potential over-recruitment budget during the first wave was 50 participants. The number of over-recruitment slots that were actually used each wave depended on the number of participants who specifically accepted the task, started the task before the current wave was complete, but did not complete the task before their wave reached the target sample size completion. For example, if six participants fell into this category during wave 1, we made six additional over-recruitment slots available during wave two to keep the total number of recruitment slots constant over time (over-recruitment slots not utilized at one wave remained available during the next wave).

This general principle of constant potential for over-recruitment was the guiding design for our recruitment strategy in both Experiments 1 and 2: it ensured fair compensation and participation opportunities for all participants, while also ensuring that no individual participant could compromise the progress of the experiment as a whole through delayed or failed participation.

Experiment 3

Participants. All experiments were approved by the institutional review board of Princeton University, and all participants gave informed consent. We recruited 960 participants from Amazon's Mechanical Turk. Participants were required to be based in the United States, have an approval rating of at least 99%, and have successfully completed 15,000 human intelligence tasks. We required higher worker metrics compared to Experiments 1 and 2 to ensure high data quality. Participants could only take the experiment once, and individuals that participated in a pilot version of Experiment 3 were excluded. Participants received a base payment of US$1.40 plus an average bonus of US$1.27. Participants spent an average of 5 min and 14 s on the experiment, for an average wage of US$35.58 per hour. Our sample size and analyses were pre-registered on 17 May 2023 (pre-registration available at ref. 80).

Conditions. All participants were assigned a motivated colour of either green or blue and received a 50 point bonus on each trial if their judgement was correct. At the end of the experiment, participants' bonuses were paid, with 10 points equal to one cent.

Participants were assigned to one of three conditions—Asocial/Motivated, Social/Correlated and Social/Resampling. Participants in the asocial condition completed the experiment on their own without observing social information. Participants in the social conditions observed a set of eight judgements made by participants in the previous wave on the same stimulus before making a judgement. These judgements were sampled with replacement from the judgements Asocial/Motivated participants in the network with the same index and previous wave made on the same stimulus. We resampled a separate set of eight judgements for each participant (for more details on resampling, see 'Oversampling and resampling').

Wave structure. Participants were recruited in two discrete waves. All participants in the first wave were assigned to the Asocial/Motivated condition, and equal numbers of participants in the second wave were assigned to each of the social conditions using block randomization. Networks and conditions were run in parallel, so participants in the second wave were not recruited until all participants in the first wave finished the experiment.

Network structure. At each wave, participants were randomly assigned to a network. We recruited 14 independent networks for each condition. A network determined the responses social participants observed. Network assignment was fixed, so a social participant observed social information from the same set of participants on each trial.

Each network had eight participants, and was composed entirely of participants with the same motivated colour. Twenty networks for each condition consisted of participants biased towards green, and 20 networks consisted of participants biased towards blue.

Procedure. We used the same stimulus task used in Experiments 1 and 2 (Fig. 1). All stimuli consisted of 100 dots displayed for one second. Participants then judged whether there were more blue or more green dots in the stimulus. A fixation cross and bounding box was displayed for 600 ms before each stimulus. Positioning and sizing for a given stimulus was determined randomly on each trial, under the constraint that no dots overlap with each other or the stimulus boundary. As in Experiment 2, we did not constrain dot positioning and sizing to match on each wave and network.

Social information on each trial was presented before participants viewed each stimulus. Social information consisted of both text and icons, with k icons coloured to match the judgement of the majority, and $8-k$ coloured in grey. For example, five green and three blue judgements were displayed as five green and three grey icons. However, if only two of eight participants chose green, participants were shown six blue icons and two grey icons. Ties were broken via simple randomization at the time of the experiment. The same social icons a participant observed before a stimulus were also presented above the blue and green response buttons when participants made their judgement.

Participants first completed two practice trials with 53 and 47 green dots. Participants received accuracy and reward feedback after both practice trials. Practice trials did not count towards participants' bonuses. Participants then completed a comprehension test before starting test rounds. Participants that failed to answer every test question correctly in their first three tries were excluded from the experiment. Test stimuli had either 48, 49, 51 or 52 green dots. Participants made judgements on 16 test stimuli, with four stimuli for every number

of green dots. Trial-by-trial feedback was not given on the test trials. Trial order for both practice and test trials was randomized.

We used the same cover stories we used in Experiments 1 and 2 to make the task more intuitive. Participants were told that they were working for an imaginary mining company looking for valuable gemstones. Participants whose motivated colour was blue were told that they were looking for blue sapphires in green grass, and those whose motivated colour was green that they were looking for green emeralds in blue water.

Oversampling and resampling. At the end of the asocial wave, we fit an IRT model to participants' judgements to estimate their bias. In these models, bias was coded in terms of bias towards green, so that participants with positive biases tend to choose green more than other participants, and those with negative biases tend to choose blue more than other participants.

In each social condition, we propagated a skewed set of judgements to participants so that asocial participants with high magnitude biases had judgements propagated at higher rates. Specifically, we assigned a probability h to propagating each participant's judgement in terms of their bias:

$$\tilde{h}_i = |b_i| \tag{15}$$

The bias weights were then normalized for all eight participants in the network:

$$h_i = \frac{\tilde{h}_i}{\sum_{k \in N} \tilde{h}_k} \tag{16}$$

On each trial, we sampled a set of eight judgements with replacement from the original set of eight judgements made by asocial participants in a given network on that stimulus. The probability of sampling each judgement was given by the participant's weight h. Participants in the Social/Correlated condition observed the skewed sampled set of judgements. Participants in the Social/Resampling condition observed a set of judgements where the skewed sample was adjusted by our resampling algorithm. To do so, we adjusted q used in equation (13) to account for oversampling:

$$q(x_{ij}) = \begin{cases} \text{logit}^{-1}(b_i - d_j) \times h_i, & \text{if } x_{ij} \text{ is green} \\ 1 - \text{logit}^{-1}(b_i - d_j) \times h_i, & \text{if } x_{ij} \text{ is blue} \end{cases} \tag{17}$$

Every other step of the resampling procedure proceeded as in Experiment 2.

The same fitted IRT model was used to estimate quantities for oversampling and resampling. We fit this model to the asocial participants' judgements after the first wave completed, running eight Markov chain Monte Carlo chains for 2,500 iterations, with 1,250 warmup iterations for each chain.

We used a bootstrapping approach to propagate judgements for both social conditions. After fitting the IRT model, we simulated running our skewing and skewing-resampling procedure 10,000 times for each judgement set (that is, each set of eight judgements for a stimulus–network combination). We used these simulations to estimate the probability of sampling a green judgement for each judgement set and used this probability to sample a set of eight judgements on each trial during the experiment. Sampling was done separately for each participant. That is, a separate sample was propagated to each social participant in the second wave.

Network simulations

Here we provide more details on the network simulations reported in the main text.

Setup. To simulate participants' judgements, we fit oracle models to the judgements made by Asocial/Motivated and Social/Motivated participants in Experiment 1. These models were highly similar to the psychometric model we fit to measure the individual-level biases of Asocial/Motivated and Social/Motivated participants in our Experiment 1 analysis (see 'Experiment 1 psychometric model' in Methods).

Both models used a Bayesian logistic regression to predict whether a participant chose green on each stimulus, and included a hierarchical term capturing the bias of each participant towards their motivated colour. The prior for the bias b_j for participant j was set to $b_j \sim \mathcal{N}(\mu_b, \sigma_b)$ with $\mu_b \sim \mathcal{N}(0, 3)$ and $\sigma_b \sim \text{lognormal}(0, 2)$. Both models also included a constant $c \sim \mathcal{N}(0, 20)$, and separate weights $\boldsymbol{y} \sim \mathcal{N}(0, 20)$ on each level of an indicator capturing the number of green dots (0.48, 0.49, 0.51 and 0.52) on the stimulus. The Social/Motivated model included an additional weight $v \sim \mathcal{N}(0, 20)$ on the number of green judgements observed for each stimulus.

We used the same 'oracle' models we used in our power analyses for Experiments 1 and 2 to simulate participants' judgements (see 'Psychometric model' in Methods; Extended Data Fig. 4). Before simulating any judgements, we first sampled a bias for each participant from the distribution defined by the mean bias parameters for the asocial oracle model. The sampled bias was multiplied by negative one for participants in blue networks, so that the bias term captured bias towards green. Participants made judgements in discrete 'iterations' (analogous to 'waves' in our experiments). On each iteration, simulated participants made judgements on a common set of stimuli. In the first iteration, participants made their decisions individually, and we used the parameters from the asocial oracle model to simulate judgements. At later iterations, participants made their judgements on the basis of the judgements of their connected peers in the network. That is, participants 'observed' the judgements made by their parents on the same stimulus in the previous iteration for every stimulus.

For each simulation, we simulated four networks; one Social/Motivated populated by participants with a motivated colour of green, one Social/Motivated network populated by participants with a motivated colour of blue, one Social/Resampling populated by participants with a motivated colour of green, and one Social/Resampling network populated by participants with a motivated colour of blue. Both the Social/Motivated and Social/Resampling networks were populated by the same set of sampled participants. Similarly, the network connections for both sets of networks were identical, and networks were 'yoked' to the same initial set of asocial decisions. That is, we first sampled a set of biases to populate both the Social/Motivated and Social/Resampling networks. We then constructed two networks (blue and green) that we used for both conditions, and used the same simulated set of asocial decisions to initialize both networks. All nodes had eight incoming edges from eight distinct nodes, and nodes were not allowed to be their own parents. Biases were sampled separately for green and blue networks, and network connections were generated for green and blue networks.

On each iteration, participants made judgements on 16 stimuli, matching the proportions we used in Experiments 2 and 3. We simulated each network for eight iterations. In the networks with Social/Resampling networks, we fit an IRT model to all participants' judgements for the previous iteration in both the blue and green networks. As in the experiment (and unlike the oracles models), this IRT model did not have access to the ground truth of any stimulus or participants' biases (see 'Experiment 2 psychometric model' in Methods). We used this IRT model to determine the resampling weights following equation (1).

We then used this model to identify weights for each judgement, and sample a new set of judgements following the same procedure we used in Experiment 2. To fit these models, we ran eight Markov chain Monte Carlo chains for 2,500 iterations, with 1,250 warmup iterations. Networks were constructed using the NetworkX[81] Python library.

Analyses. We limited our analyses to comparing bias and accuracy on iterations where simulated participants observed social information—that is, we ignored data from the first iteration. To test for differences between conditions, we fit two logistic regression models on the social data from both conditions and the relevant dependent variable (chose motivated colour or chose correct); one with a coefficient for condition, and one without. Both models included a fixed effect for motivated colour (participant and network random intercepts were excluded as they led to singular fits; including them does not change any of our findings). We then determined whether bias or accuracy varied between the conditions by performing a likelihood ratio test between the two models.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Experiment and simulation data for this study are available through the Open Science Repository at https://doi.org/10.17605/OSF.IO/YTH5R.

Code availability

Code for both experiments, data analyses and simulations are available at https://doi.org/10.17605/OSF.IO/YTH5R, which contains an archived version of a GitHub repository containing all of the experiment code. Both experiments were built using `Dallinger` (Experiment 1: 5.1.0; Experiment 2: custom fork; Experiment 3: 9.0.0). The resampling algorithm in Experiments 2 and 3 were implemented in Python (3.9) using `NumPyro` (0.7.2). The resampling algorithm for the power analyses and the psychometric models of Experiment 1 were implemented in R using `Rstan` (2.21.3). The predictions shown in Extended Data Fig. 1 were generated using MATLAB. Mixed effects models were implemented in R (4.1.2) using `lme4` (1.1.28). Network simulations were done in Python (3.10.4) using the NetworkX library (3.1).

References

1. Lerman, K. & Ghosh, R. Information contagion: an empirical study of the spread of news on digg and twitter social networks. In *Fourth International AAAI Conference on Weblogs and Social Media* (AAAI Press, 2010).
2. Bakshy, E., Rosenn, I., Marlow, C. & Adamic, L. The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web* 519–528 (ACM, 2012).
3. Lerman, K. Social information processing in news aggregation. *IEEE Internet Comput.* **11**, 16–28 (2007).
4. Hermida, A. in *The SAGE Handbook of Digital Journalism* (eds Witschge, T. et al.) 81–94 (SAGE Publications, 2016).
5. Gainous, J. & Wagner, K. M.*Tweeting to Power: The Social Media Revolution in American Politics* (Oxford Univ. Press, 2013).
6. Burns, K. S. *Celeb 2.0: How Social Media Foster Our Fascination with Popular Culture* (ABC-CLIO, 2009).
7. Bakshy, E., Messing, S. & Adamic, L. A. Exposure to ideologically diverse news and opinion on Facebook. *Science* **348**, 1130–1132 (2015).
8. Conover, M. et al. Political polarization on Twitter. *Proc. International AAAI Conference on Web and Social Media* **5**, 89–96 (2011).
9. Pariser, E. *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think* (Penguin, 2011).
10. Levy, R. Social media, news consumption, and polarization: evidence from a field experiment. *Am. Econ. Rev.* **111**, 831–70 (2021).
11. Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W. & Starnini, M. The echo chamber effect on social media. *Proc. Natl. Acad. Sci. USA* **118**, e2023301118 (2021).
12. Shin, J. & Thorson, K. Partisan selective sharing: the biased diffusion of fact-checking messages on social media. *J. Commun.* **67**, 233–255 (2017).
13. Settle, J. E. *Frenemies: How Social Media Polarizes America* (Cambridge Univ. Press, 2018).
14. Allcott, H., Gentzkow, M. & Yu, C. Trends in the diffusion of misinformation on social media. *Res. Politics* **6**, 2053168019848554 (2019).
15. Allcott, H. & Gentzkow, M. Social media and fake news in the 2016 election. *J. Econ. Perspect.* **31**, 211–36 (2017).
16. Tucker, J. A. et al. Social media, political polarization, and political disinformation: a review of the scientific literature. *SSRN Electron. J.* https://doi.org/10.2139/ssrn.3144139 (2018).
17. Yarchi, M., Baden, C. & Kligler-Vilenchik, N. Political polarization on the digital sphere: a cross-platform, over-time analysis of interactional, positional, and affective polarization on social media. *Polit. Commun.* **38**, 98–139 (2021).
18. Brady, W. J., McLoughlin, K., Doan, T. N. & Crockett, M. J. How social learning amplifies moral outrage expression in online social networks. *Sci. Adv.* **7**, eabe5641 (2021).
19. Eom, Y.-H. & Jo, H.-H. Generalized friendship paradox in complex networks: the case of scientific collaboration. *Sci. Rep.* **4**, 1–6 (2014).
20. Alipourfard, N., Nettasinghe, B., Abeliuk, A., Krishnamurthy, V. & Lerman, K. Friendship paradox biases perceptions in directed networks. *Nat. Commun.* **11**, 707 (2020).
21. Jackson, M. O. The friendship paradox and systematic biases in perceptions and social norms. *J. Polit. Econ.* **127**, 777–818 (2019).
22. Balkin, J. M. How to regulate (and not regulate) social media. *Knight Institute Occasional Paper Series* https://knightcolumbia.org/content/how-to-regulate-and-not-regulate-social-media (2020).
23. Cusumano, M., Gawer, A. & Yoffie, D. Social media companies should self-regulate. Now. *Harvard Business Review* https://hbr.org/2021/01/social-media-companies-should-self-regulate-now (2021).
24. Lazer, D. M. et al. The science of fake news. *Science* **359**, 1094–1096 (2018).
25. Bail, C. et al. Social-media reform is flying blind. *Nature* **603**, 766 (2022).
26. Mason, W. & Watts, D. J. Collaborative learning in networks. *Proc. Natl Acad. Sci. USA* **109**, 764–769 (2012).
27. Becker, J., Brackbill, D. & Centola, D. Network dynamics of social influence in the wisdom of crowds. *Proc. Natl Acad. Sci. USA* **114**, E5070–E5076 (2017).
28. Jayles, B. et al. How social information can improve estimation accuracy in human groups. *Proc. Natl Acad. Sci. USA* **114**, 12620–12625 (2017).
29. Rendell, L. et al. Why copy others? Insights from the social learning strategies tournament. *Science* **328**, 208–213 (2010).
30. Smaldino, P. E. & Richerson, P. J. Human cumulative cultural evolution as a form of distributed computation. in *Handbook of Human Computation* (ed Michelucci, P.) 979–992 (Springer, 2013).
31. Moussaïd, M., Brighton, H. & Gaissmaier, W. The amplification of risk in experimental diffusion chains. *Proc. Natl Acad. Sci. USA* **112**, 5631–5636 (2015).
32. Luo, M., Hancock, J. T. & Markowitz, D. M. Credibility perceptions and detection accuracy of fake news headlines on social media: effects of truth-bias and endorsement cues. *Commun. Res.* **49**, 171–195 (2022).
33. Kahneman, D., Slovic, S. P., Slovic, P. & Tversky, A. *Judgment under Uncertainty: Heuristics and Biases* (Cambridge Univ. Press, 1982).
34. Hastorf, A. H. & Cantril, H. They saw a game; a case study. *J. Abnorm. Soc. Psychol.* **49**, 129–134 (1954).

35. Dunning, D. & Balcetis, E. Wishful seeing: how preferences shape visual perception. *Curr. Dir. Psychol. Sci.* **22**, 33–37 (2013).
36. Leong, Y. C., Hughes, B. L., Wang, Y. & Zaki, J. Neurocomputational mechanisms underlying motivated seeing. *Nat. Hum. Behav.* https://doi.org/10.1038/s41562-019-0637-z (2019).
37. Bruner, J. S. & Goodman, C. C. Value and need as organizing factors in perception. *J. Abnorm. Soc. Psychol.* **42**, 33–44 (1947).
38. Balcetis, E. & Dunning, D. See what you want to see: motivational influences on visual perception. *J. Person. Soc. Psychol.* **91**, 612–625 (2006).
39. Griffiths, T. L. & Kalish, M. L. Language evolution by iterated learning with Bayesian agents. *Cogn. Sci.* **31**, 441–480 (2007).
40. Rendell, L., Fogarty, L. & Laland, K. N. Rogers' paradox recast and resolved: population structure and the evolution of social learning strategies. *Evolution* **64**, 534–548 (2010).
41. Mesoudi, A. & Whiten, A. The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philos. Trans. R. Soc. B* **363**, 3489–3501 (2008).
42. Miton, H. & Charbonneau, M. Cumulative culture in the laboratory: methodological and theoretical challenges. *Proc. R. Soc. B* **285**, 20180677 (2018).
43. Boyd, R. & Richerson, P. J. *Culture and the Evolutionary Process* (Univ. Chicago Press, 1988).
44. Henrich, J. *The Secret of Our Success: How Culture is Driving Human Evolution, Domesticating Our Species, and Making Us Smarter* (Princeton Univ. Press, 2016).
45. Laland, K. *Darwin's Unfinished Symphony* (Princeton Univ. Press, 2017).
46. Hardy, M. D., Krafft, P. M., Thompson, B. & Griffiths, T. L. Overcoming individual limitations through distributed computation: rational information accumulation in multigenerational populations. *Top. Cogn. Sci.* **14**, 550–573 (2022).
47. Thompson, B., Van Opheusden, B., Sumers, T. & Griffiths, T. Complex cognitive algorithms preserved by selective social learning in experimental populations. *Science* **376**, 95–98 (2022).
48. Olsson, A., Knapska, E. & Lindström, B. The neural and computational systems of social learning. *Nat. Rev. Neurosci.* **21**, 197–212 (2020).
49. Acemoglu, D., Dahleh, M. A., Lobel, I. & Ozdaglar, A. Bayesian learning in social networks. *Rev. Econ. Stud.* **78**, 1201–1236 (2011).
50. Liu, R. & Xu, F. Learning about others and learning from others: Bayesian probabilistic models of intuitive psychology and social learning. *Adv. Child Dev. Behav.* **63**, 309–343 (2022).
51. Embretson, S. E. & Reise, S. P. *Item Response Theory for Psychologists* (Psychology Press, 2013).
52. Balietti, S., Getoor, L., Goldstein, D. G. & Watts, D. J. Reducing opinion polarization: effects of exposure to similar people with differing political views. *Proc. Natl Acad. Sci. USA* **118**, e2112552118 (2021).
53. Guilbeault, D., Becker, J. & Centola, D. Social learning and partisan bias in the interpretation of climate trends. *Proc. Natl Acad. Sci. USA* **115**, 9714–9719 (2018).
54. Bail, C. A. et al. Exposure to opposing views on social media can increase political polarization. *Proc. Natl Acad. Sci. USA* **115**, 9216–9221 (2018).
55. Tokdar, S. T. & Kass, R. E. Importance sampling: a review. *Wiley Interdisc. Rev. Comput. Stat.* **2**, 54–60 (2010).
56. Elvira, V., Martino, L., Luengo, D. & Bugallo, M. F. Generalized multiple importance sampling. *Stat. Sci.* **34**, 129–155 (2019).
57. Mitchell, A., Gottfried, J., Kiley, J. & Matsa, K. E. Political polarization & media habits. *Pew Research Center's Journalism Project* https://policycommons.net/artifacts/619536/political-polarization-media-habits/1600676 (2014).
58. Centola, D. & Macy, M. Complex contagions and the weakness of long ties. *Am. J. Sociol.* **113**, 702–734 (2007).
59. Granovetter, M. S. The strength of weak ties. *Am. J. Sociol.* **78**, 1360–1380 (1973).
60. Watts, D. J. *Small Worlds: The Dynamics of Networks between Order and Randomness* (Princeton Univ. Press, 2004).
61. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998).
62. Centola, D., Eguíluz, V. M. & Macy, M. W. Cascade dynamics of complex propagation. *Physica A* **374**, 449–456 (2007).
63. Watts, D. J. The "new" science of networks. *Annu. Rev. Sociol.* **30**, 243–270 (2004).
64. Feld, S. L. Why your friends have more friends than you do. *Am. J. Sociol.* **96**, 1464–1477 (1991).
65. Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P. & Bhattacharjee, B. Measurement and analysis of online social networks. In *Proc. of the 7th ACM SIGCOMM Conference on Internet Measurement* 29–42 (ACM, 2007).
66. Kumar, R., Novak, J. & Tomkins, A. Structure and evolution of online social networks. In *Proc. of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 611–617 (ACM, 2006).
67. Dunbar, R. I., Arnaboldi, V., Conti, M. & Passarella, A. The structure of online social networks mirrors those in the offline world. *Soc. Netw.* **43**, 39–47 (2015).
68. Whalen, A., Griffiths, T. L. & Buchsbaum, D. Sensitivity to shared information in social learning. *Cogn. Sci.* **42**, 168–187 (2018).
69. Molavi, P., Tahbaz-Salehi, A. & Jadbabaie, A. Foundations of non-Bayesian social learning. *Columbia Business School* https://ssrn.com/abstract=2683607 (2017).
70. DeVito, M. A. From editors to algorithms: a values-based approach to understanding story selection in the facebook news feed. *Digit. Journal.* **5**, 753–773 (2017).
71. Lazer, D. The rise of the social algorithm. *Science* **348**, 1090–1091 (2015).
72. Walter, N., Cohen, J., Holbert, R. L. & Morag, Y. Fact-checking: a meta-analysis of what works and for whom. *Polit. Commun.* **37**, 350–375 (2020).
73. Guess, A. M. et al. A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proc. Natl Acad. Sci. USA* **117**, 15536–15545 (2020).
74. Grimmelmann, J. The virtues of moderation. *Yale J. Law Technol.* **17**, 42–109 (2015).
75. Arganda, S., Pérez-Escudero, A. & de Polavieja, G. G. A common rule for decision making in animal collectives across species. *Proc. Natl Acad. Sci. USA* **109**, 20508–20513 (2012).
76. https://osf.io/yth5r
77. Hoffman, M. D. & Gelman, A. The no-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15**, 1593–1623 (2014).
78. Chmielewski, M. & Kucker, S. C. An MTurk crisis? Shifts in data quality and the impact on study results. *Soc. Psychol. Person. Sci.* **11**, 464–473 (2020).
79. https://osf.io/87me6
80. https://osf.io/8s7y2
81. Hagberg, A., Swart, P. & Schult, D. *Exploring Network Structure, Dynamics, and Function using NetworkX* (Los Alamos National Lab, 2008).

Acknowledgements

This work was made possible with funding T.L.G. received from the NOMIS Foundation. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the paper.

Author contributions

All authors contributed to designing the experiment, developing the mitigation algorithm, and writing the paper. M.D.H. and B.D.T. implemented and ran the experiments and simulations. M.D.H. analysed the results.

Competing interests

T.L.G. has previously received research funding from Facebook/Meta, a social media company. This funding was for a separate research project and has not supported this research. The authors declare no other competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41562-023-01715-5.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41562-023-01715-5.

Correspondence and requests for materials should be addressed to Mathew D. Hardy.

Peer review information *Nature Human Behaviour* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

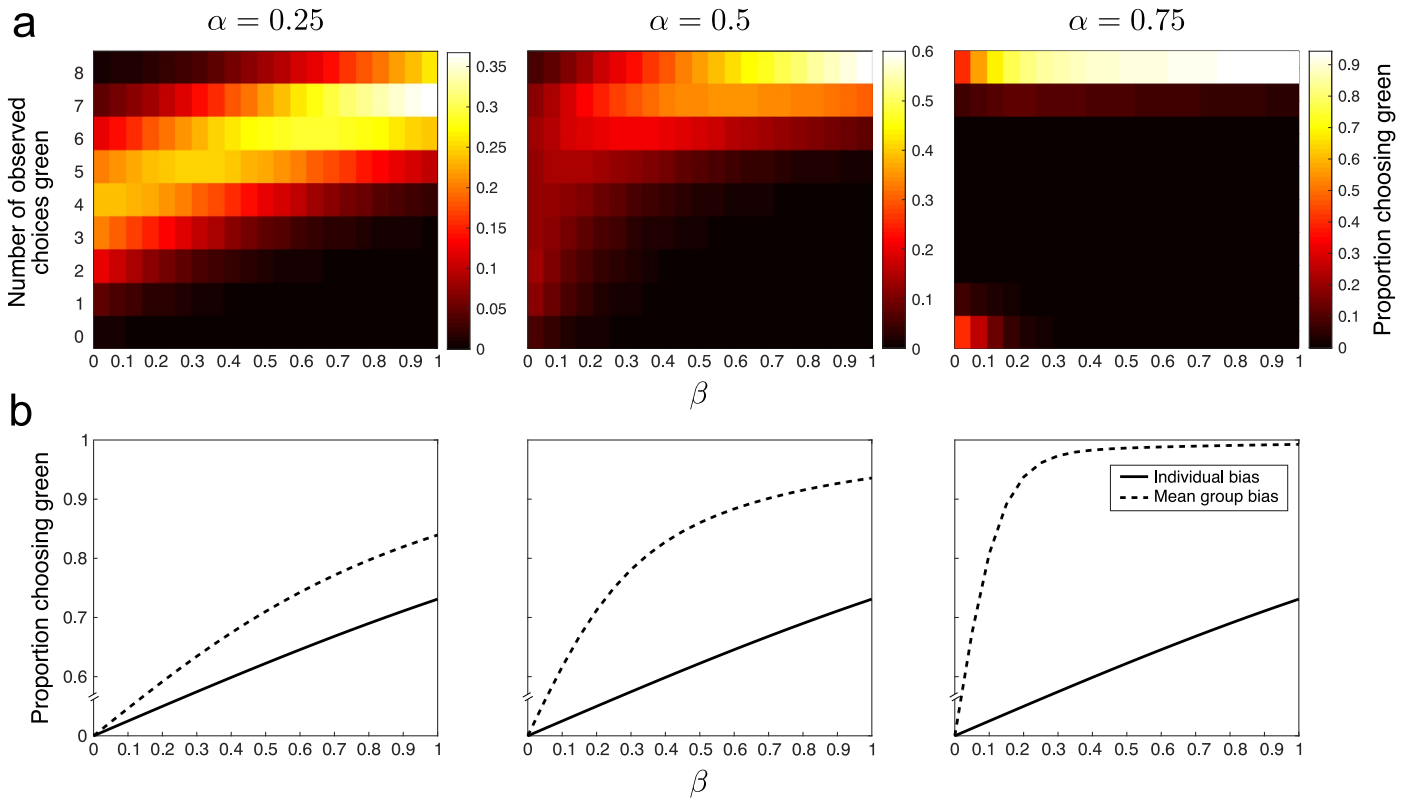
Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

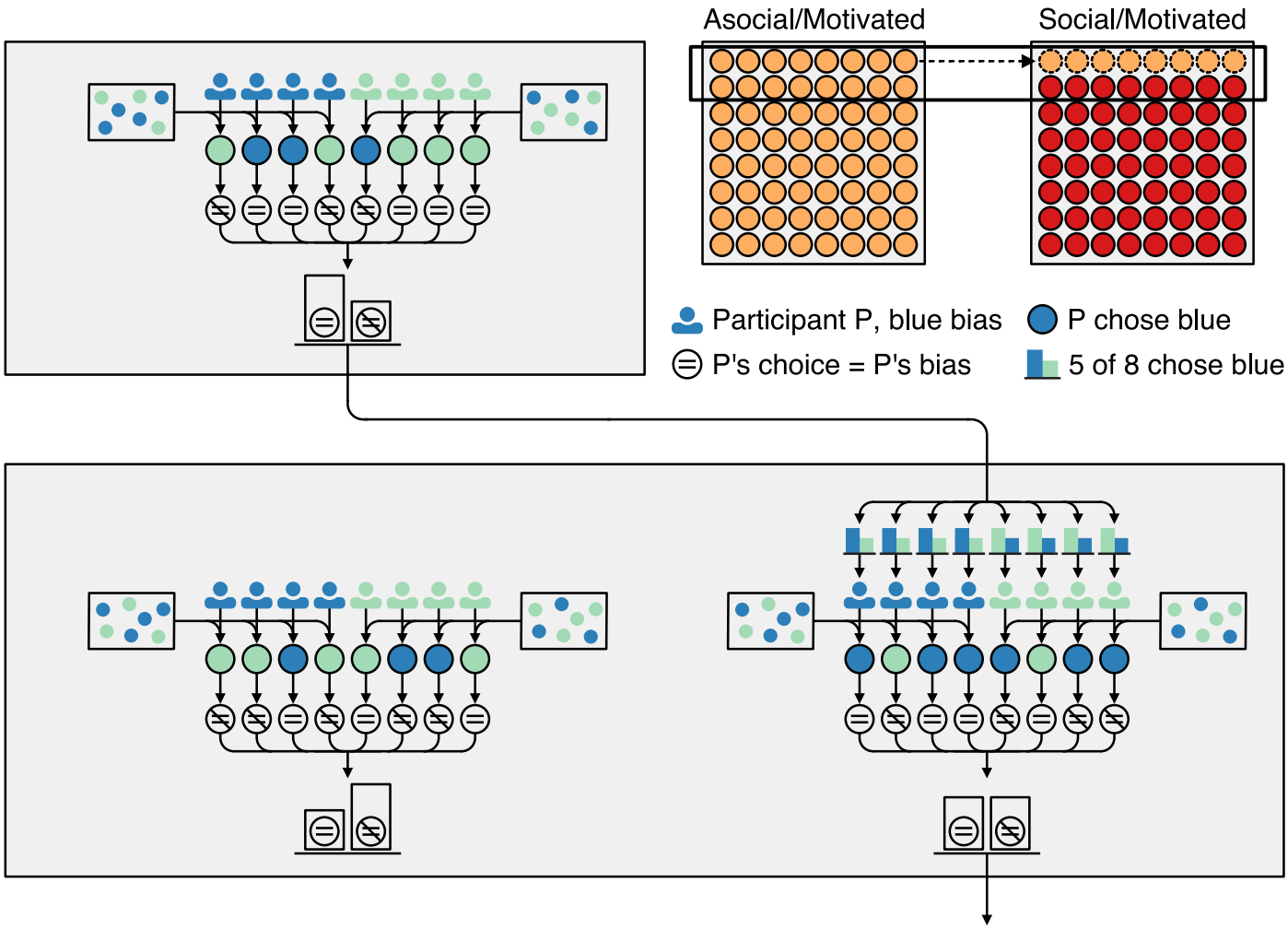
© The Author(s), under exclusive licence to Springer Nature Limited 2023

https://doi.org/10.1038/s41562-023-01715-5



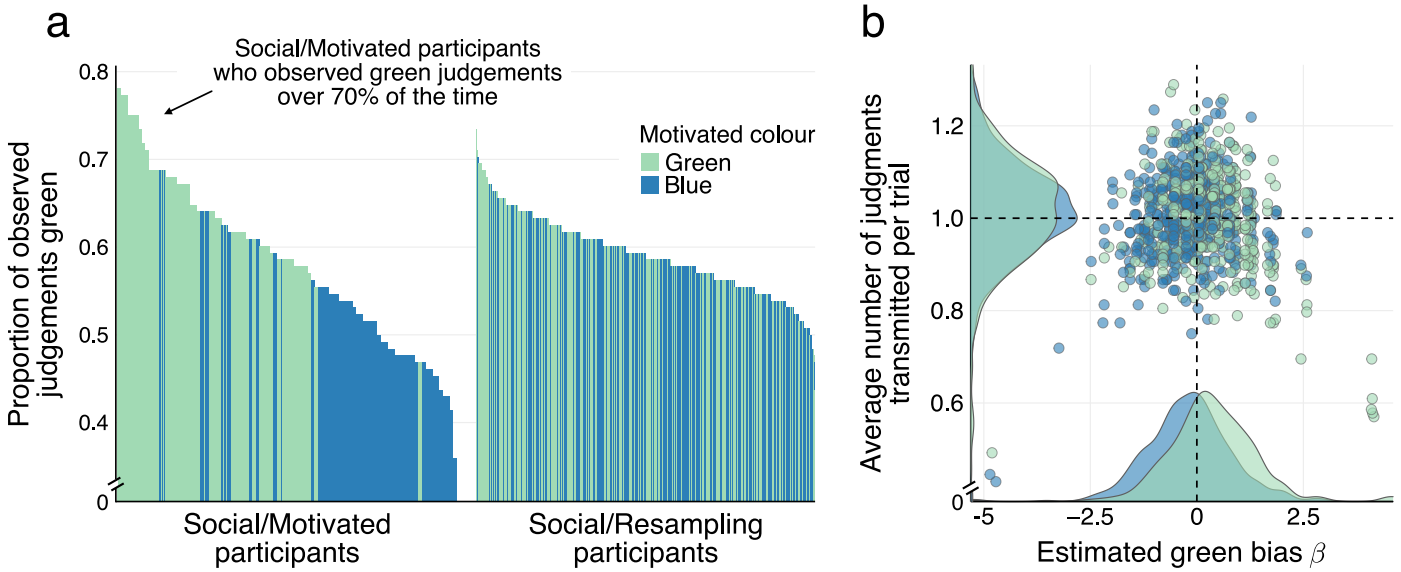
Extended Data Fig. 1 | Stationary distributions for social networks. (a) Stationary distributions on the proportion of people endorsing green as a function of the bias β towards green, for different levels of sensitivity to social information α. The bias translates into a stationary distribution strongly skewed towards green, with increasing effect as α increases. (b) Average proportion of green judgements under this stationary distribution, compared against the bias of a single individual. The social network amplifies individual biases. Because bias β and the information from the stimulus γ_d have the same effect in the mathematical equation on judgements, this model predicts that the effects of both will be exaggerated by participating in a social network: people will become more biased, but also more accurate. Note that this analysis could be equivalently stated in terms of blue bias and blue judgements.

https://doi.org/10.1038/s41562-023-01715-5



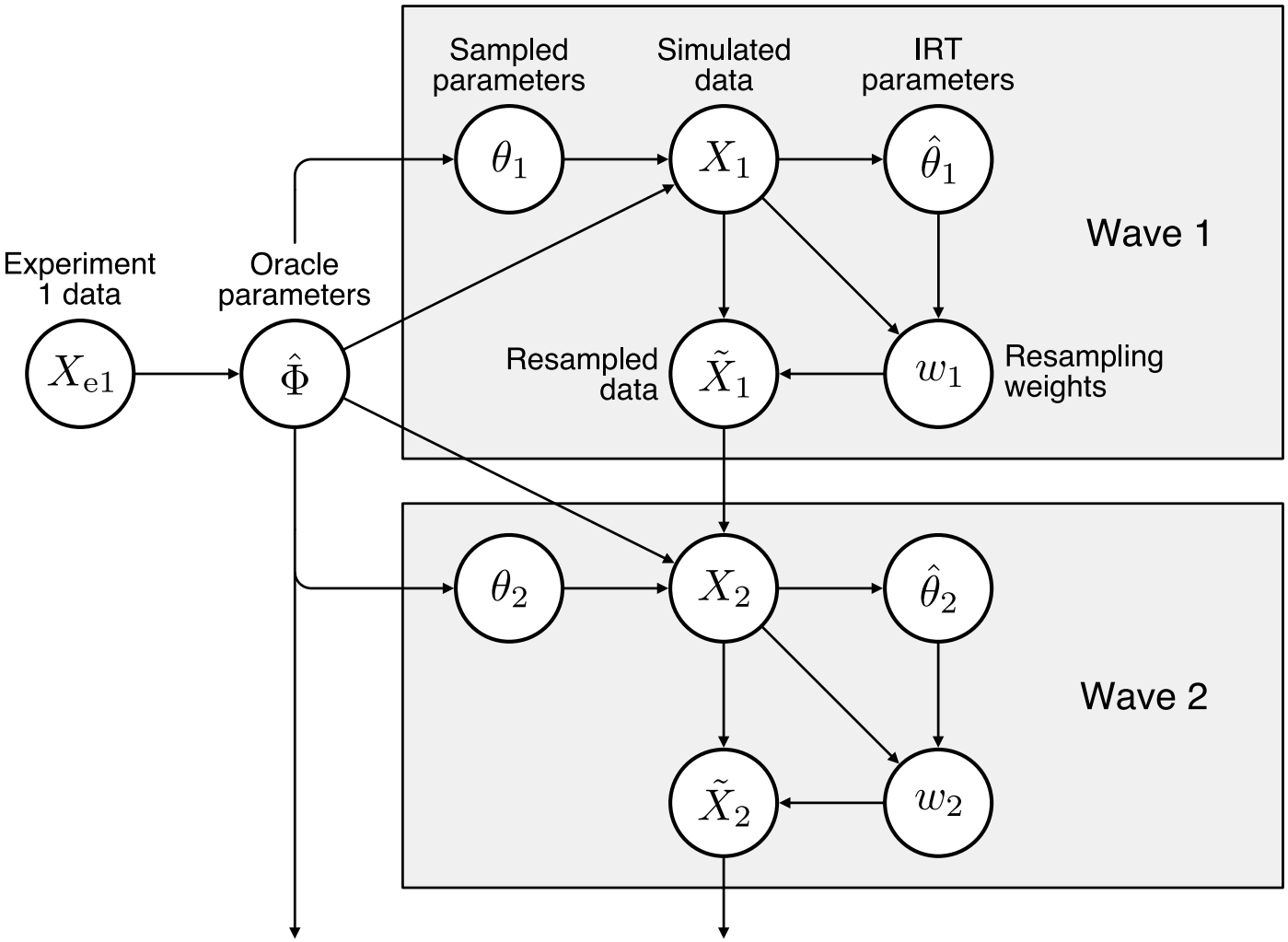
Extended Data Fig. 2 | Experiment 1 yoking structure and marked colour design. All participants in the first wave were assigned to an asocial condition. In the second wave, social participants observed the judgements made by first-wave asocial participants in the same motivated condition and network index (see Methods, Experiment 1). Each network consisted of four participants with a marked colour of blue and four of green at each wave. A participant's marked colour determined the colour of the dots in the stimulus (dot sizes and positions were fixed in a wave and network). To match this colour-swapping, social information was also presented in terms of the participant's marked colour. That is, if 5 of 8 people chose their marked colour in the previous wave, participants with a marked colour of green would be told that 5 of 8 people chose green, and participants with a marked colour of blue that 5 of 8 participants chose blue. Marked colour corresponded to motivated colour for participants in motivated conditions. Participants in neutral conditions were assigned marked colours using an identical process, but were not informed of their marked colour.

https://doi.org/10.1038/s41562-023-01715-5



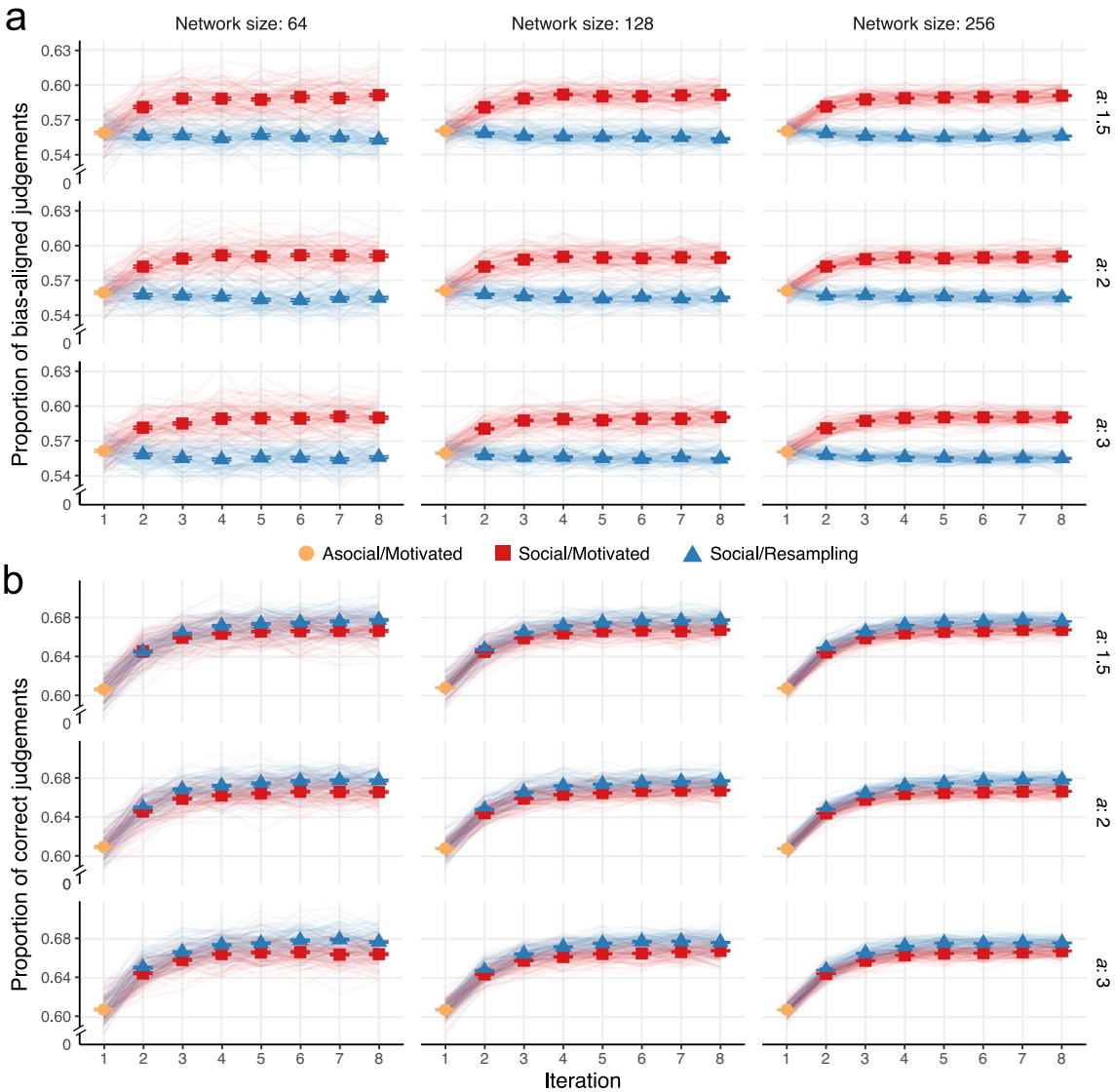
Extended Data Fig. 3 | Experiment 2 participant observations and estimated biases. The resampling algorithm increased consensus between networks with induced biases towards green and blue. (a) Each bar shows the proportion of green judgements from the previous wave that each participant (n=784 for both conditions) observed over all 16 trials of the experiment. Bars are arranged in descending order, and bar colour corresponds to the participant's motivated colour. (b) Estimated participant biases and transmission rates in the Social/Resampling condition. In our resampling algorithm, each participant's judgement could be propagated multiple times to a participant at the next wave. Rather than only propagating judgements made by those with low estimated bias, the algorithm transmitted each participant's judgements at similar rates. Points show the estimated green biases of participants in the Asocial/Motivated (wave 1) and Social/Motivated (waves 2-7) condition and the number of times their judgements were transmitted.

https://doi.org/10.1038/s41562-023-01715-5



Extended Data Fig. 4 | Simulation design. For the network simulations and power analyses, we alternated between simulating participants' judgements and the effects of our resampling procedure. We first fit the parameters Φ of the oracle models to Experiment 1 data X_{e1} using Markov Chain Monte Carlo. One model was fit to Asocial/Motivated participants, and one to Social/Motivated participants. For the power analyses, at each wave t, we used either the asocial (wave 1) or social (waves 2-8) oracle to sample participant biases θ_t and simulate judgements θ_t (this process is illustrated here). By contrast, for the network simulations, we sampled a set of participant biases θ using the social oracle model and used these parameters to simulate judgements X_t at each iteration (there was no population turnover in our network simulations). To simulate our resampling procedure, we then fit IRT parameters θ_t to the simulated judgements at each wave. As in Experiments 1 and 2, these IRT models did not have access to the ground truth or participants' true biases. We used our fitted IRT model to determine the importance weights w_t for each judgement and resample a set of judgements $\tilde{X}_t$ to propagate to the next wave or iteration. For each simulation, we repeated this process for 8 waves (the same fitted oracle model was used in all simulations).

https://doi.org/10.1038/s41562-023-01715-5



Extended Data Fig. 5 | Network simulation results. Plots show the proportion of (a) bias-aligned and (b) correct judgements by iteration. Simulated participants in the first iteration made their judgements asocially. For each set of simulations, two histories were sampled starting at the second iteration; one where participants viewed the judgements made by their parents in the previous iteration (Social/Neutral), and one where these judgements were resampled using our algorithm (Social/Resampling). The same constructed networks were used for both histories, with one populated by simulated participant "paid" for blue dots, and one by simulated participant "paid" for green. Error bars show the standard errors of the proportions (n=204,800, 409600, and 819,200 for each point in the networks with 64, 128, and 256 participants, respectively).

https://doi.org/10.1038/s41562-023-01715-5

Extended Data Table 1 | Experiment 1 results

		Asocial/Neutral						
		Reported	Prereg					
	Bias	13.4	12.7					
Asocial/ Motivated	p	$< .001$	$< .001$					
	Accuracy	0.1	0.1	**Asocial/Motivated**				
	p	.75	.75	Reported	Pre-reg			
	Bias	1.7	1.6	23.5	22.3			
Social/ Neutral	p	.19	.20	$< .001$	$< .001$			
	Accuracy	25.6	25.6	22.6	22.6	**Social/Neutral**		
	p	$< .001$	$< .001$	$< .001$	$< .001$	Reported	Prereg	
	Bias	83.4	78.5	31.5	29.9	102.4	96	
Social/ Motivated	p	$< .001$	$< .001$	$< .001$	$< .001$	$< .001$	$< .001$	
	Accuracy	17.6	17.6	15.1	15.1	0.7	0.7	
	p	$< .001$	$< .001$	$< .001$	$< .001$	0.40	0.40	**Social/Motivated**
Mean bias		0.526		0.562		0.512		0.618
Mean accuracy		0.604		0.608		0.654		0.646

The values in each 2×4 cell give chi-square statistics and p-values for the likelihood ratio tests between two logistic regressions comparing the relevant conditions. The dependent variable in bias comparisons captures whether a participant chose their marked colour (see Methods, Experiment 1 for a description of marked colour), and whether they chose correctly for accuracy comparisons. All models include random intercepts for each network to control for our dot randomization scheme (see Methods, Experiment 1). Reported and prereg models are identical, except that prereg models include random intercepts for each participant. Note that the statistics from these two models will be nearly identical when the variance of the participant random intercepts is near zero. The bottom two rows give the proportion of judgements where participants chose their marked colour (bias) and the correct colour (accuracy) for each condition. p-values were not adjusted for multiple comparisons.

https://doi.org/10.1038/s41562-023-01715-5

Extended Data Table 2 | Parameter estimates from the psychometric models fit to Experiment 1

	Asocial / Motivated	Social / Motivated
Bias mean	0.268 [0.218, 0.32] 1.23	0.371 [0.302, 0.44] 1.00
Bias standard deviation	0.069 [0.017, 0.155] 1.81	0.462 [0.353, 0.566] 1.02
48 green dots indicator	−0.449 [−0.546, −0.358] 1.03	−1.409 [−1.557, −1.2563] 1.00
49 green dots indicator	−0.24 [−0.336, −0.146] 1.02	−1.068 [−1.227, −0.909] 1.00
51 green dots indicator	0.416 [0.32, 0.505] 1.02	−0.41 [−0.603, −0.223] 1.00
52 green dots indicator	0.704 [0.606, 0.815] 1.03	−0.316 [−0.525, −0.106] 1.00
Green endorsements		0.232 [0.199, 0.265] 1.00
Number of samples	10,000	10,000

Each estimate gives the mean sampled value for the parameter averaged over all chains and iterations. Values in brackets below each estimate show cutoffs for the 90% credible intervals. Values below the brackets show Rhat values for the statistic. Running the asocial model for more iterations results in lower values of Rhat but has negligible impacts on parameter estimates. Individual-level participant biases are not reported.

https://doi.org/10.1038/s41562-023-01715-5

Extended Data Table 3 | Experiment 2 results

		Asocial/Motivated				
		Reported	Pre-reg			
Social/ Motivated	Bias	12.5	7.6			
	p	$< .001$	.006			
	Accuracy	82.5	26.1	**Social/Motivated**		
	p	$< .001$	$< .001$	Reported	Pre-reg	
Social/ Resampling	Bias	0.5	3.8	7.5	11.3	
	p	.48	.05	.006	$< .001$	
	Accuracy	59.3	10.5	1.9	1.9	
	p	$< .001$	.001	.17	.17	**Social/Resampling**
Mean bias		0.54		0.567		0.545
Mean accuracy		0.574		0.63		0.621

As in Extended Data Table 1, the values in each 2×4 cell give chi-square statistics and p-values for the likelihood ratio tests between logistic regression models with and without separate fixed effects for the two conditions. Bias models predicted whether participants chose their motivated colour, and accuracy models whether they chose correctly. All models include a fixed effect for motivated colour, and a random intercept for each participant. Prereg models include an additional random intercept for each social network to account for the social yoking scheme (see Methods, Experiment 2). The bottom two rows give the proportion of judgements where participants chose their motivated colour (bias) and the correct colour (accuracy) for each condition. p-values were not adjusted for multiple comparisons.

https://doi.org/10.1038/s41562-023-01715-5

Extended Data Table 4 | Experiment 3 results

		Asocial/Motivated		Social/Correlated		Social/Resampling
		Reported	Pre-reg			
Social/ Correlated	Bias	14.7	14.9			
	p	$< .001$	$< .001$			
	Accuracy	1.0	1.1			
	p	.31	.30			
Social/ Resampling	Bias	1.0	9.7	8.0	17.3	
	p	.31	.002	.005	$< .001$	
	Accuracy	2.7	2.7	0.4	0.4	
	p	.10	.10	.54	.53	
Mean bias		0.527		0.576		0.540
Mean accuracy		0.584		0.594		0.599

The values in each 2×4 cell give chi-square statistics and p-values for the likelihood ratio tests between two logistic regression models; one with separate fixed effects for the two conditions, and one where both conditions were coded as a single condition. Bias models predict whether participants chose their motivated colour, and accuracy models whether they made the correct judgement. All models include a fixed effect for motivated colour and a random intercept for each participant. Prereg models have an additional random intercept for each social network to account for the social yoking scheme (see Methods, Experiment 3; we reported the "Reported" models as the prereg models had singular fits). The bottom two rows give the proportion of judgements where participants chose their motivated colour (bias) and the correct colour (accuracy) for each condition. p-values were not adjusted for multiple comparisons.

https://doi.org/10.1038/s41562-023-01715-5

Extended Data Table 5 | Network simulation results

			a: 1.5	a: 2.0	a: 3.0
Network size: 64	Bias	Social/Motivated	0.588	0.590	0.588
		Social/Resampling	0.555	0.555	0.556
		$\chi^2(1)$	3158.7	3512.3	3107.5
		p	$< .001$	$< .001$	$< .001$
	Accuracy	Social/Motivated	0.662	0.661	0.661
		Social/Resampling	0.669	0.671	0.672
		$\chi^2(1)$	162.7	307.6	349.2
		p	$< .001$	$< .001$	$< .001$
Network size: 128	Bias	Social/Motivated	0.589	0.588	0.588
		Social/Resampling	0.555	0.555	0.556
		$\chi^2(1)$	6705.2	6377.8	6041.7
		p	$< .001$	$< .001$	$< .001$
	Accuracy	Social/Motivated	0.662	0.662	0.661
		Social/Resampling	0.670	0.670	0.670
		$\chi^2(1)$	402.8	396.0	523.9
		p	$< .001$	$< .001$	$< .001$
Network size: 256	Bias	Social/Motivated	0.588	0.589	0.589
		Social/Resampling	.556	0.556	0.556
		$\chi^2(1)$	12387.4	12575.7	12583.7
		p	$< .001$	$< .001$	$< .001$
	Accuracy	Social/Motivated	0.662	0.661	0.661
		Social/Resampling	0.670	0.670	0.670
		$\chi^2(1)$	843.0	1063.9	907.5
		p	$< .001$	$< .001$	$< .001$

Values report the average proportions of bias-aligned judgements, correct judgements, and chi-square statistics testing for differences between the resampling and neutral conditions (values were averaged across all 100 simulations and iterations 2-8). All models include a fixed effect for motivated colour (random intercepts for each participant and simulation number were excluded as they lead to singular fits). Chi-square tests were done by running a likelihood ratio test between an unrestricted model with fixed effects for both conditions (resampling and motivated) and a restricted model without this fixed effect. p-values were not adjusted for multiple comparisons.

nature portfolio

Corresponding author(s): Mathew Hardy

Last updated by author(s): 08/26/2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All code and data is available at https://doi.org/10.17605/OSF.IO/YTH5R. Experiments were built using Dallinger (Experiment 1: 5.1.0; Experiment 2: 7.7.0; Experiment 3: 9.0.0). The resampling algorithm in Experiment 2 was implemented in Python (3.9) using NumPyro (0.7.2). The resampling algorithm for the power analyses for Experiments 2 and 3 and the psychometric models of Experiment 1 were implemented in R using Rstan (2.21.3). Predictions from the Bayesian model shown in Figure 2 were generated using MATLAB. Network simulations were done in Python (3.10.4) using the NetworkX library.
Data analysis	Mixed effects models were implemented in R (4.1.2) using lme4 (1.1.28).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Experiment and simulation data for this study are available through the Open Science Repository at https://doi.org/10.17605/OSF.IO/YTH5R.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Data on participants' sex and gender was not collected.
Population characteristics	Data on participants' demographics was not collected.
Recruitment	Participants for all studies were recruited online through Amazon Mechanical Turk. Our sample was therefore potentially not fully representative. Participants were randomly assigned to treatments.
Ethics oversight	All experiments were approved by the institutional review board of Princeton University.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	All data are quantitative. On each round of the experiment, participants observed randomly positioned and sized blue and green dots for one second and estimated the color of the majority of dots. Our experimental design varied the presence of an induced bias (control and motivated conditions) and social information (asocial and social conditions). All participants received a reward on each trial if their judgement was correct. Participants in motivated conditions received an additional bonus on every trial for every dot of their motivated colour. Participants in asocial conditions made their choices individually, whereas participants in social conditions observed the choices made by other participants on the same stimulus.
Research sample	All participants were recruited on Amazon Mechanical Turk and so samples were not representative. Participants were required to be located in the United States of America and at least 18 years old. We limited participants to those living in the US as our IRB only covered US participants. For Experiment 1, we required a minimum Mechanical Turk approval rating of 95. For Experiment 2, we required participants to have a minimum Mechanical Turk approval rating of 95 and 10,000 approved HITs. For Experiment 3, we required a Mechanical Turk approval rating of 99 and 15,000 approved HITs. We had stricter requirements for Experiments 2 and 3 to keep up with evolving data quality norms (i.e., to reduce the proportion of bot participants).
Sampling strategy	We preregistered our sample size before data collection for all experiments. For Experiment 1, we selected the largest sample size our budget permitted. We selected the sample size for Experiments 2 and 3 using a power analysis on simulated data of our proposed experiment (see Supplementary Information for more information). All participants were recruited from the MTurk participant pool, and so all samples were convenience samples.
Data collection	For all experiments, participants were recruited in discrete waves (batches) over several days. Participant recruitment was done using the Dallinger Software package. To speed data collection, we recruited participants using an algorithm that partially recruited participants at each generation. This allowed us to speed data collection given the constraint that all participants in a generation had to complete the experiment before the next generation was recruited. Recruitment was done via Mechanical Turk, and so participants completed the experiment using their own computer. No individual besides the participant was required to be present. Participants were not informed of which condition they were assigned to nor the hypotheses being tested.

Timing	Data for Experiment 1 was collected in September 2019. Data for Experiment 2 was collected in November 2021. Data for Experiment 3 was collected in May 2023.
Data exclusions	For Experiments 2 and 3, our recruitment algorithm partially overrecruitted participants at each generation to speed data collection. As specified in our preregistrations, we excluded data from overrecruited participants from analysis. We will also exclude a participant's data from analysis if the experiment webpage fails for that participant.
Non-participation	We excluded participants who did not complete the full experiment, who failed to pass a comprehension quiz in their first three tries, or who were over-recruited (see preregistrations for more details). In Experiment 1, we excluded 219 of 2,619 recruited participants. In Experiment 2, we excluded 447 of 2911 recruited participants. In Experiment 3, we excluded data from 157 from 1,117 recruited participants.
Randomization	Participants were assigned to a treatment using block randomization.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging