



OPEN

## Rational inference strategies and the genesis of polarization and extremism

Peter D. Kvam<sup>1</sup>✉, Abhay Alaukik<sup>1</sup>, Callie E. Mims<sup>2</sup>, Arina Martemyanova<sup>1</sup> & Matthew Baldwin<sup>1</sup>

Polarization and extremism are often viewed as the product of psychological biases or social influences, yet they still occur in the absence of any bias or irrational thinking. We show that individual decision-makers implementing optimal dynamic decision strategies will become polarized, forming extreme views relative to the true information in their environment by virtue of how they sample new information. Extreme evidence enables decision makers to stop considering new information, whereas weak or moderate evidence is unlikely to trigger a decision and is thus under-sampled. We show that this information polarization effect arises empirically across choice domains including politically-charged, affect-rich and affect-poor, and simple perceptual decisions. However, this effect can be disincentivized by asking participants to make a judgment about the difference between two options (estimation) rather than deciding. We experimentally test this intervention by manipulating participants' inference goals (decision vs inference) in an information sampling task. We show that participants in the estimation condition collect more information, hold less extreme views, and are less polarized than those in the decision condition. Estimation goals therefore offer a theoretically-motivated intervention that could be used to alleviate polarization and extremism in situations where people traditionally intend to decide.

Attitudes toward crucial economic, environmental, and social policies are becoming increasingly polarized<sup>1,2</sup>. Extreme, and polarized views that are exacerbated by social and motivational forces can lead to civil unrest and violence<sup>3</sup>. Meanwhile, polarization and extremism have begun to manifest more frequently beyond the political sphere, extending to divisions and conflict among juries, boardrooms, religious organizations, nonpartisan commissions<sup>4</sup>, sports, and even consumer product brands (Apple vs Google, car brands, or the “console wars”)<sup>5</sup>. Even in entirely mundane decisions about perceptual stimuli without any political charge, opinions can become polarized<sup>6</sup>. We often have little choice in the matter: Individuals' efforts and intentions to avoid forming extreme or polarized views can be rendered ineffective in the face of social and persuasive processes that exacerbate differences in opinion among groups<sup>7</sup>.

Polarization and extremism may not be so concerning if there were an easy solution to reversing their effects on cognition and behavior. Once it has set in, polarization distorts our view of the world and the way we consider new information, causing us to seek out confirming evidence<sup>8–10</sup> both because confirming information is pleasurable and because it is felt to be more valuable than disconfirming information<sup>11</sup>. Managing the dissonance that arises from exposure to counter-attitudinal information leads to greater polarization, as decision makers with opposing attitudes begin to deal in extremes. Common ground is lost as information is selected not to inform but rather to reinforce negative attitudes toward one's opponents<sup>12,13</sup>. In concert with these confirmation biases, people also strive to disconfirm information that is not in line with their existing beliefs<sup>14</sup>. In terms of reasoning, this can be perfectly rational: Bayesian reasoning prescribes that we should not ignore our prior beliefs when forming new ones, even when processing counter-attitudinal information<sup>15</sup>. However, prior beliefs or expectations can affect how people interpret new information, resulting in them taking more extreme positions or forming more extreme beliefs in light of information that conflicts with their existing beliefs<sup>14,16–18</sup>. People engage in motivated reasoning and belief revision driven by a desire to form beliefs that align with others<sup>19</sup> and the utilities associated with holding and maintaining those beliefs<sup>20</sup>. This is exacerbated by the patterns of behavior on social networks, where communication and network dynamics lead to increases in polarization and the creation of echo chambers among users who agree with one another<sup>21–23</sup>. Attempts to reverse the process of polarization can even backfire<sup>4,16,17</sup>, as experimental findings have shown that groups of US Twitter users on opposite sides of the political spectrum became *more* polarized when exposed to opposing political views<sup>24</sup>.

<sup>1</sup>University of Florida, Gainesville, USA. <sup>2</sup>University of South Alabama, Mobile, USA. ✉email: pkvam@ufl.edu

Considering how insidious polarization can be once it has taken root, it is critical to understand and prevent the process of polarization at its source. In this paper, we seek to explain how individual-level cognitive mechanisms can create polarization in the first place, examining how people gather information as they consider new situations. Specifically, we evaluate how *inference goals* interact with decisions about when to continue or stop sampling information when choosing among options<sup>25,26</sup>, and how different goals can create or ameliorate polarization at the genesis of new opinions or views. Computational modeling of these choice processes has revealed that individual-level decision dynamics are related to political beliefs and ideological preferences across a range of topics like authoritarianism and dogmatism, social dominance, and openness to opposing views more generally<sup>27</sup>. Here, we go a step further to show that the very process of decision-making, even executed by a perfectly rational decision-maker, will create polarization and extreme views relative to the true information in the decision-maker's environment. That is, rational individual decision-makers who try to make accurate and efficient choices will still become polarized and adopt extremist views. And although we build on work suggesting that rational goals can create polarization in social communication and group environments<sup>28,29</sup>, observing polarization in our findings would suggest that rational individuals can form extreme and polarized views even in absence of communication or social pressure, in purely individual decisions, about relatively neutral and novel information ecologies.

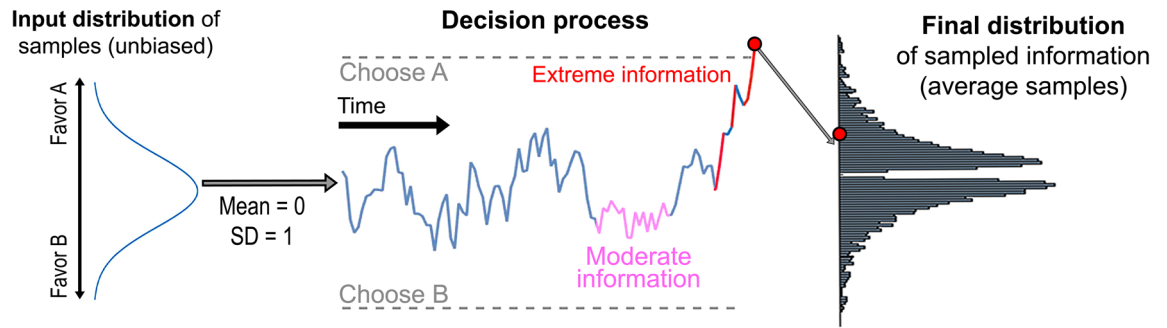
If polarization is inevitable when making decisions, how can it be stopped? Fortunately, making a decision is not the only goal people can have when gathering new information. As opposed to determining which of two options is superior, one can evaluate or estimate their relative merit. Such a strategy incentivizes *precision* rather than *discrimination*, changing the conditions under which people stop sampling new information<sup>26</sup>. As a result, changing the incentives or directions for a task to favor estimation may prevent people from forming polarized views in the first place. This shifts the target of polarization interventions to one of prevention, rather than treatment: Estimation incentives should keep people from ever becoming polarized or forming extreme views (relative to the true information “out there” in the world) in the first place. In the following sections, we outline the reasoning behind this kind of intervention and how it avoids polarization and extremism caused by decision goals, briefly note the results of a re-analysis of seven decision-making studies illustrating the prevalence of polarization, and introduce a new study where we manipulate the goals of information collection to reduce the extremeness and polarization of views people form.

**The rational decision maker.** Prior work often paints a picture of polarization as a phenomenon that arises from biases in information processing or belief updating<sup>30</sup>, motivation<sup>8,9</sup>, ingroup-outgroup processes<sup>31,32</sup>, and media algorithms and design<sup>33</sup>. Each of these explanations for polarization assumes that polarization emerges from bias or irrationality in the way decisions are made, while others explain how these beliefs can be propagated or exacerbated by rational reasoning based on already-polarized prior beliefs<sup>14,16,17,27,34–37</sup>. While such biases or priors certainly perpetuate divisions between people and exacerbate the problem of extremism, they do not entirely explain the genesis of polarization and how similar people can wind up on opposite sides of an innocuous issue at the outset. In this paper, we suggest that polarization can be created on entirely novel issues from basic judgment and decision making processes themselves. Moreover, we argue that these processes can act independently of social forces, such as motivation or group-processes, and can lead to polarization even in the absence of initial biases or social interaction. To support this claim, we prove mathematically that decision-makers will be biased toward extreme views when implementing optimal choice strategies aimed at maximizing decision quality and efficiency. This proof is provided in its completeness in the supplementary material, but we summarize it here.

When a decision maker is presented with a choice between two (or more) options, they must make a compromise between investing time on that decision and making a choice so that they can move on. Both time and belief accuracy bestow utility—the utility of making good (accurate) decisions is clear, but making a quick decision allows a decision-maker to proceed to the next choice, or move on to a task that they enjoy more than the choice. If a decision maker focuses purely on collecting information, they will never be able to stop and make a choice, whereas if they focus entirely on decision speed then they will have at-chance accuracy and make generally poor choices. Under the general framework of random walk/diffusion processes that are commonly used to describe and explain choice behavior in humans and other animals<sup>38–40</sup>, the optimal trade-off is achieved by setting a pre-defined *threshold* that corresponds to the level of accuracy that the decision maker desires, and then gathering information until that threshold is reached<sup>41,42</sup>. These models are based on the sequential probability ratio test<sup>42,43</sup>, where a decision maker tracks the balance of support between two options and uses this balance to determine when to make a choice. As they get information favoring option A, they shift up toward the “Choose A” response boundary, as shown in Fig. 1. Similarly, as they get information favoring option B, they shift down toward the “Choose B” response boundary. Once they cross one boundary or the other, the corresponding alternative is chosen (see the supplement for a full mathematical proof and description of simulations).

This strategy is specifically optimal for maximizing *reward rate*, or the amount of rewards that a decision-maker can accumulate per unit of time (seconds, minutes, days). There are slight modifications to the strategy, such as collapsing choice boundaries<sup>44,45</sup>, that are made to adapt it to different environments, but all reward rate optimality models share the elements of (a) tracking the balance of support, (b) accumulating the relative balance to a threshold, and then (c) deciding when the balance crosses a boundary. However, as we illustrate below, these three elements lead decision-makers to a distorted view of the evidence in their environment despite the decision process being optimal.

This unintended outcome occurs because a decision-maker's goal is not to gather a representative sample of information, but to tip the balance of information far enough in one direction to conclude  $A > B$  or  $B > A$  (where  $>$  denotes a preference order or belief). That is, choice tasks will bias information samples in the direction



**Figure 1.** Illustration of how an unbiased distribution of information (left) becomes a polarized distribution (right) when fed through the decision process (middle). Because of an incentive to choose between options, moderate information—since it does not give credence to either options—gets discarded (middle).

of the choice. On one hand, this is not so surprising, because we should expect people to have choice-consistent samples: Those who sample more of choice A information should choose option A. However, a deeper analysis reveals a surprising corollary to this expected pattern, namely, that the information sampled up to the point of choice is not representative of the true information in the world. Rather than mirroring the information in the environment (about choice A for instance), a decision-maker's sampled information is more extreme<sup>25,26</sup>. In other words, the *goals* of the decision-makers influence the final sample of information that they collect, creating an unrepresentative sample of information despite no inherent biases.

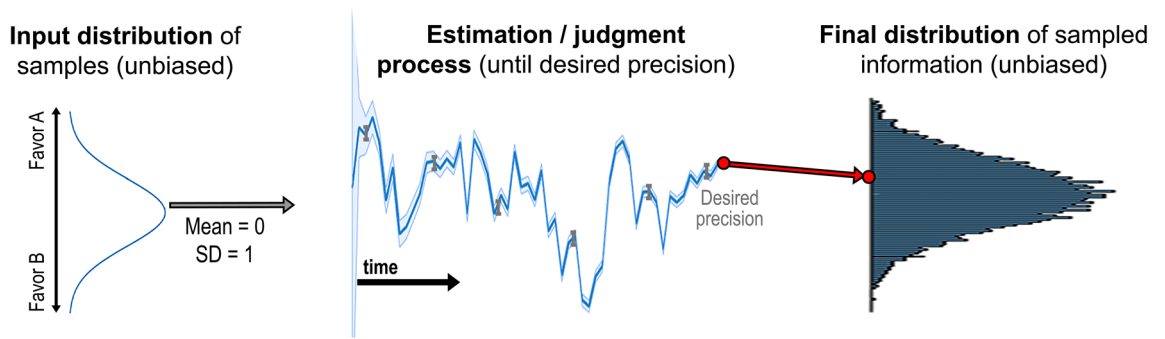
One explanation for why this occurs is because moderate information is not conducive to achieving an agent's decision goals (to stop and support  $A > B$ , or  $B > A$ ), while extreme information makes it easy to stop and select one option over the other. These ideas are illustrated in Fig. 1: Moderate information (pink) cannot tip the balance of evidence sufficiently toward one option or another to trigger a choice, while extreme information (red) is very likely to do so. As a result, people who have made a decision using a balance-of-evidence stopping rule will have a greater number of extreme information samples than the underlying distribution.

In addition to polarization, optimal decision strategies can create subgroups of highly polarized individuals whose information samples are at the extreme ends of the spectrum. When a person gathers a piece of strong information in their first sample, there is the chance that this is the only piece of information they need to consider to make a choice. As a result, some individuals are likely to construct their views and make their decisions based on a minimal sample of particularly extreme information. In the simulations in the supplementary material, we examine factors that exacerbate the problem of these “extremists” who gather very little information and yet hold much more extreme views than other decision makers. In general, these extremist decision makers tend to be implementing optimal strategies that place greater subjective value on time over new information relative to other decision makers, reflected in lower thresholds that generate fast but not necessarily accurate responses<sup>46,47</sup>.

In the supplementary material, we provide a re-analysis of seven existing studies that allowed participants to make decisions by sampling information piece by piece over time. These studies show that the information participants consider is more extreme than the information naturally provided by their environment, lending credence to our prediction that decision-makers will become polarized. Furthermore, these studies show that the participants with the most extreme views are the same participants who sample the smallest amount of information, showing that extremism can arise even on simple perceptual and preferential choices. But the emergence of polarization and extremism is not inevitable. In the next section we describe one way that the decision process can be manipulated to incentivize the sampling of more representative information from the world.

**Interventions to reduce polarization.** The seven re-analyzed studies illustrate that polarization and extremism both arise frequently around even mundane choices (numbers of dots, which posters/foods are best), and it may seem inevitable in all decisions. Fortunately, decision-making is not the only goal that people can pursue when making inferences. Strategies that are non-optimal for reward rate maximization, but satisfy other criteria, can be useful for reducing polarization among groups of people. For example, a strategy elicited by incentivizing estimation accuracy—where a person is asked to assess *how much* the information in their environment favors one option relative to another—can lead participants to gather a distribution of information that is representative of the true information in the environment.

Some support for the efficacy of an estimation goal in reducing the sampling issues in decision making is provided by Coenen and Gureckis<sup>26</sup>. They found that participants gathered largely representative samples while estimating relative strength of information, as opposed to a bimodal distribution of samples while choosing between two options. Model-based simulation results provide further evidence for its efficacy. As illustrated in Fig. 2, an optimal strategy for making an estimate (minimizing response time for a desired level of error) requires a participant (i.e., the judge) to continue sampling new information until a desired level of precision is achieved. This process involves collecting new information and shifting their estimate over time (blue line) until the error of their estimate of the mean (shaded blue area can correspond to standard error or posterior variance of the mean) falls below a criterion level. In this case, extreme information is counterproductive to making precise estimates because it fails to reduce the variance of a sample, and thus the posterior variance of the mean. Conversely, information that is close to the current posterior mean results in maximal reduction of the posterior variance



**Figure 2.** Illustration of a precision-based judgment process, where the goal is not to select one side over the other (as in Fig. 1) but to meet a criterion level of estimation precision.

of the mean estimate. As a result, the judge is incentivized to gather more information to form an estimate after gathering extreme information, and incentivized to stop after gathering moderate information.

Following this optimal strategy, the distribution of sampled information across a population of decision makers (right side of Fig. 2) will correspond closely to the input distribution (left side of Fig. 2). Posing the task as an estimation problem can therefore be used as an effective manipulation that reduces polarization and extremism by bringing the evidence sample closer to the true information distribution. In this paper, we test the effectiveness of this intervention across multiple types of inference problems, encompassing both simple perceptual problems and more politically or affectively charged situations. In the next section, we introduce an experiment that examines these possibilities and explores the individual differences that drive or exacerbate polarization and extremism in some people.

## Methods

All methods were carried out in accordance with relevant ethical guidelines and regulations. All experiments involving primary data were approved as exempt by the University of Florida Institutional Review Board (IRB202002176). These experiments were deemed to involve minimal risk to participants. Prior to completing the experiments, all participants were briefed on the study procedures and completed informed consent.

The experiments we present here were collected in two parts. An initial study collected data from 40 participants in each condition, separately for the choice and estimation conditions. We then collected data from a preregistered replication, randomly assigning an additional 110 participants to either the choice or estimation condition. However, the behavioral data from these two stages of data collection were nearly identical. We therefore present them altogether here for the sake of brevity. The data files from the first stage of data collection (Stage 1) and from the preregistered replication (All Data) are both provided on the Open Science Framework at [osf.io/jsb52](https://osf.io/jsb52), so readers can view and analyze the two data sets separately as they wish. The preregistration is available at [osf.io/qfb6w](https://osf.io/qfb6w).

Our initial re-analysis showed that choice tasks, by their very structure, lead to polarized and extreme information sampling, even when applying rational and optimal sampling strategies. Moreover, our own simulations as well as work by previous authors<sup>26</sup> suggest that estimation tasks could serve as a solution to this problem by reducing the extent to which people make decisions on the basis of extreme information. We further elucidate these processes in a new study that allows us to test whether polarization arises naturally in new environments with politically and affectively charged decisions. By comparing decision and estimation conditions, we evaluate whether estimation is a strategy that can alleviate polarization and extremism. The new study allows us to formalize our definitions of polarization and extremism (outlined below), and examine them as individual differences that can be related to relevant psychological traits, such as dogmatism, need for structure, and other demographic variables<sup>27</sup>. In the sections to follow, we detail each of these measures and examine their relationship to information sampling and individual-level metrics of polarization and extremism.

The structure of choice and estimation tasks were matched closely: In both conditions, corresponding scenarios were worded almost identically, the layout of the window was kept identical, and the sampling action (clicking on a button to collect information about Policy A vs. B) was kept the same as well. The key difference between the two conditions was in the way participants responded to each scenario. In the choice condition, they were asked to decide between two courses of action (“Select Policy A or Policy B”). In the estimation condition, they were asked to estimate the relative expected effect of implementing the two courses of action on a relevant criterion (rate how good is Policy A vs Policy B). This allowed us to directly compare patterns of information sampled in the choice and estimation conditions for every question that we asked.

**Task and items.** In both conditions, participants were first instructed to imagine that they had encountered an alien civilization seeking their advice on several of their planet’s pressing issues. The use of an “alien world” cover story is designed to minimize the influence of biases that participants might bring to each scenario and to make it clear to participants that they are gathering information with the goal of forming beliefs about an entirely new issue. The alien worlds paradigm has been commonly used to examine processes related to the evaluative information ecology<sup>48</sup>, acquisition of morality<sup>49</sup>, language learning<sup>50</sup>, and reinforcement learning<sup>51</sup>.

The aliens are building a new power plant that will generate energy from one of two materials - Material A and Material B. Your task is to gather information about these materials using your equipment and recommend which material they should choose.

**Measure Energy Output**  
 Measurements show that material A generates about 38 kWh MORE energy than material B.  
 Measurements show that material A generates about 11 kWh MORE energy than material B.  
 Measurements show that material A generates about 12 kWh LESS energy than material B.

Once you are ready to choose, select one of these options:

Recommend Material A      Recommend Material B

The aliens are building a new power plant that will generate energy from one of two materials - Material A and Material B. Your task is to gather information about these materials using your equipment and estimate the difference in output between Material A and Material B.

**Measure Energy Output**  
 Measurements show that material A generates about 6 kWh MORE energy than material B.  
 Measurements show that material A generates about 2 kWh MORE energy than material B.  
 Measurements show that material A generates about 30 kWh LESS energy than material B.

Once you are ready, enter your estimate using the scale below:

Material A will generate 50kWh LESS than Material B      Material A will generate 50 kWh MORE than Material B

-50   -40   -30   -20   -10   0   10   20   30   40   50

Estimated difference in energy output:

**Figure 3.** The scenarios in both tasks are worded very closely. The key difference lies in the task: after gathering information by clicking the button, participants either choose one of two options (left), or estimate how much better one option is than the other (right; the scale ranges from  $-2.5$  to  $+2.5$  SD of the true information distribution).

Category	Representative example
Affect-rich	Pursue gene therapy or immunotherapy to save aliens from a deadly virus
Affect-poor	Mining for the mineral Xenon in Location A or B
Politically charged	Allowing immigrant aliens into the city or not
Risky	Settling in city A or B based on earthquake frequency
Investment	Investing in flood prevention infrastructure now or spend money later on flood response

**Table 1.** Representative examples of issues under each category of scenario.

Participants were informed that the aliens were considering two possibilities for each issue (i.e., “The aliens are building a new power plant that will generate energy from one of two materials—Material A and Material B”) and participants were tasked with helping the aliens decide by sampling information about each of the options. Participants had access to a button that would generate the relevant information (e.g., “Each unit of Material A produces  $X$  MORE/LESS units of energy than each unit of Material B”). Each time participants clicked the button, the random variable  $X$  was drawn from a normal distribution with mean 0,  $X \sim N(0, SD)$ . This meant that the true information that participants could sample was normally distributed and (on average) did not favor one option over the other. Participants had the freedom to sample as much or as little as they wished before responding.

Crucially, participants in the choice condition were asked to choose the option which they thought was better (i.e., Material A or B), whereas participants in the estimation condition were asked to assess the degree to which one option was better than the other using a sliding scale (Fig. 3) that encouraged participants to view the scenario as distinct from a binary choice. The actual choice or estimate that participants made was largely unimportant to the analyses we present below—instead, our aim was to evaluate the relative degree of polarization and extremism between the choice and estimation conditions by comparing patterns of information search in each scenario.

Participants in both conditions experienced a total of 80 scenarios: 40 in which the button enabled them to sample information from the world (i.e. gathering information using their “equipment” or “tools”) and 40 in which the button allowed for sampling information from others (i.e. participants could consult with local alien experts to obtain relevant information). The scenarios were designed to incorporate a variety of issues that differed in terms of their subject matter as well as the degree to which they evoked affective involvement<sup>52,53</sup>: the 80 questions included 20 affect-poor, 16 affect-rich, 22 politically charged, 8 risky, 10 investment, and 4 attention-check scenarios. A complete list of the choice scenarios, along with the data and analyses presented here, is provided on the Open Science Framework at [osf.io/jsb52](https://osf.io/jsb52). Representative examples of scenarios under each category are shown in Table 1.

**Participants.** We recruited a total of 95 Prolific Academic workers for each condition for a total of 180 participants. This was composed of two stages of data collection: an initial sample of 40 participants in the choice and estimation conditions, assigned to conditions based on when they signed up to participate; and a preregistered replication study that randomly assigned 110 participants to either the choice or estimation condition.

Our goal was to obtain data from approximately 85 participants in each condition, which is sufficient to identify relationships between individual-difference measures when they are present and to support the null hypothesis with a Bayes factor of 3 or more (moderate evidence) when there is no relationship between measures. A formal power analysis is provided in the supplement.

Ten participants from the choice condition and eleven from the estimation condition were removed because they failed attention checks or failed to respond at all on multiple trials. Hence, we retained responses from 85 individuals in the choice condition and 84 in the estimation condition, very close to our goal of 85 participants per condition. Participants in the choice condition were 54 women (30 men, 1 non-binary/other) with a mean age of 33.62 ( $SD = 11.63$ ). For the estimation condition, participants were 45 women (37 men, 2 non-binary/other) with an average age of 36.59 ( $SD = 12.71$ ). All participants indicated that they were from the United States and fluent English speakers. Participants were paid \$10 per hour for their participation in the experiment.

Our sample size was intended to allow us 80% “power”—in this case, the ability to conclude in favor of the null using a Savage-Dickey Bayes factor analysis<sup>54</sup> when there was no relationship between variables in a regression. Moreover, we collected responses to 80 trials for each participant, which results in approximately 13,500 data points in total to estimate the effects. Such a large set of data helps make our parameter estimates of both group-level and individual-level characteristics robust and reliable. Moreover, most of our analyses involve hierarchical (mixed effect) Bayesian models, which are typically stricter than classical statistics under uninformative priors<sup>55</sup>. In the results, we also provide 95% credible/highest density intervals [HDIs] for our point estimates, which indicate the 95% posterior most likely estimates of a particular parameter or statistic.

**Individual difference analyses.** Similar to the re-analysis of the seven empirical studies in the previous section, we compared the observed distribution of information collected by participants against the expected distribution of an unbiased, representative sample that we would expect from a random sample of information. These distributions were compared in two main ways. First, we computed the difference between the distributions using an information-theoretic metric, the Kullback-Leibler divergence [KLD]. This quantifies the difference between the expected probability density of samples (a unimodal, normal distribution) and the observed density of samples. It is used as our measure of polarization, as it quantifies the degree to which the samples gathered by participants diverge from an expected unimodal distribution.

Second, we also quantify the extremeness of the sampled information by evaluating its variance, divided by the true variance of the pool of information participants are sampling from (i.e., the variance of a representative sample). This allows us to assess the degree to which participants’ sampling strategies led them to draw samples that are more (or less) extreme than the true information provided by their environment. Note that this measure of extremeness is often correlated with, but distinct from, our measure of polarization. As shown in Fig. 4, a distribution of information can be highly polarized/bimodal without high variance, and likewise it can have high variance without being bimodal or polarized.

Putting it all together, we formally define polarization and extremism in these studies as follows:

- **Polarization** Polarization is the extent of divergence between the distribution of the information collected and the true (expected) information distribution. We operationalize it as the Kullback-Liebler Divergence [KLD], which provides a measure of the overall difference between two probability distributions that we can use to compare observed and expected information distributions:

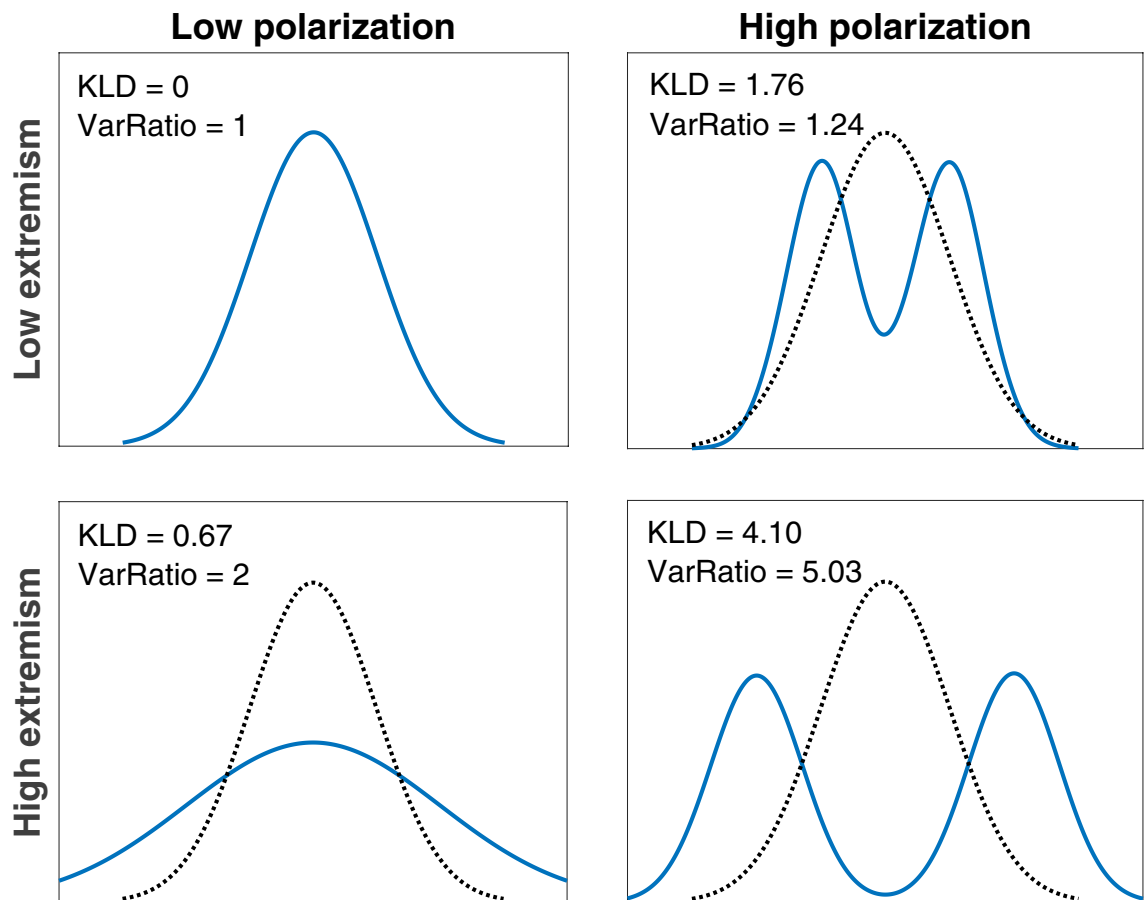
$$KLD[p(y)||p(x)] = \int_i^N p(y_i) \log_n \frac{p(y_i)}{p(x_i)} \quad (1)$$

where  $p(y_i)$  describes the probability density of a standard normal distribution and  $p(x_i)$  is a nonparametrically estimated probability density, generated using a Gaussian kernel density estimator with optimal bandwidth<sup>56</sup>, of sampled data. The indices  $i : N$  represent each of points on the probability distributions where their divergence is calculated.

- **Extremism** The ratio of the variance of the observed sample to the variance of a representative sample of information from that trial—termed *variance ratio* for the purposes of this study. If the overall variance ratio in the condition (i.e., condition-level) is above 1, the information sampled in the condition will be deemed extreme.

Both of these metrics can be applied at the level of a trial, participant, or entire condition: We simply evaluate the KLD or variance of a single trial relative to the expected distribution on that trial, all trials from a participant relative to the expected distribution for that participant, or all trials from all participants in a condition, respectively. In all cases, we compare the observed to expected distributions *conditional on how many samples were drawn*. This allows us to disentangle the number of samples drawn from the shape of the distribution of samples they drew. It also allows us to examine the number of samples as a separate outcome variable. The number of samples (the number of clicks on the “draw another sample” button in Fig. 3) served as a simple and direct measure of how much information people chose to gather before making their choice or estimate. We also measured the length of time participants spent on each trial in seconds, although this data was not analyzed.

For the extremism analysis/variance ratio, we can estimate uncertainty in the variance parameter simply by fitting it using Bayesian MCMC sampling methods. However, the KLD is a point estimate and thus does not convey uncertainty, making it hard to directly compare across conditions. To estimate a 95% HDI on the KLD for decision and estimation tasks, we instead simulated 10,000 artificial data sets with the same characteristics (number of trials, number of samples per trial) as the observed data. However, these artificial data sets were



**Figure 4.** Diagram of the measures of polarization (low = left, high = right) and extremism (low = top, high = bottom), how different distributions of information might exemplify each of these properties, and how KLD and variance ratio (VarRatio) describe/quantify these properties. The dotted black line corresponds to the expected distribution, a normal distribution centered at zero, while the blue line corresponds to the observed distribution with high or low polarization/extremeness.

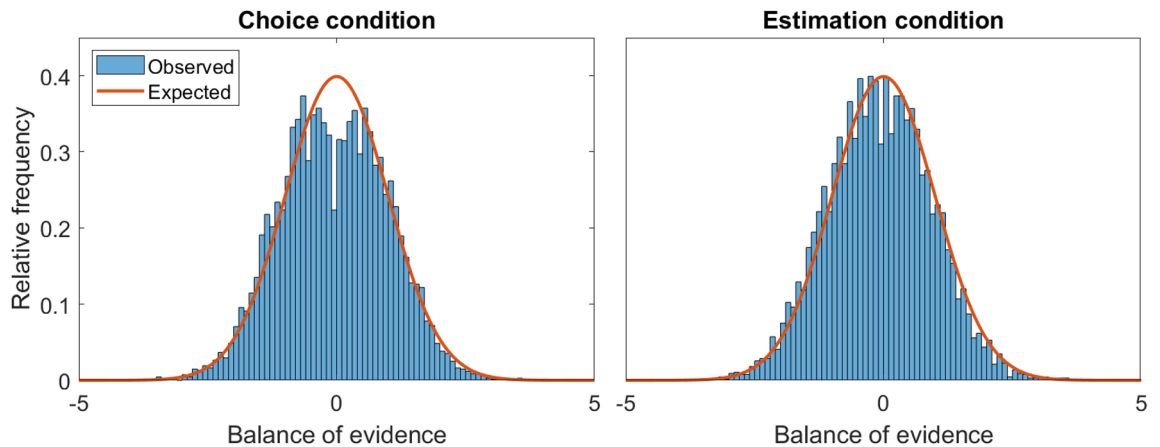
sampled from the true underlying distribution of information from the stimulus, i.e., they did not use the same stopping rules as participants. In other words, they did not have any sampling biases that our participants might have had. This yielded 10,000 “samples” from the true distribution against which we could compare the observed results. We computed the mean difference between observed and expected KLDs by subtracting the mean of the KLDs across all 10,000 artificial data sets from the KLD of the observed data set. Then, we constructed the 95% HDI on the difference between the observed and expected distribution by computing the 95% highest density interval of differences between observed KLD and each of the 10,000 artificial data set KLDs.

At the conclusion of the study, participants completed self-report measures of dogmatism<sup>57</sup> and personal need for structure<sup>58</sup> and reported their political ideology and basic demographic information. We included dogmatism in our analysis because previous work has connected it to reduced information sampling under uncertainty<sup>59</sup>. Need for structure was included because need for closure, of which the need for structure appears to be a more psychometrically desirable variant<sup>58</sup>, has previously been connected to the threshold parameter in dynamic decision models<sup>60</sup>. We report the relationships between all of these individual differences—extremism/variance ratio, polarization/KLD, information sampling/number of clicks, dogmatism, and need for structure—in the results.

## Results

All the analyses were conducted in R<sup>61</sup>, JASP<sup>62</sup>, and MATLAB. Data were analyzed with Bayesian estimation of the posterior using uninformed priors and hierarchical/mixed models, where appropriate. We implemented MCMC sampling in JAGS<sup>63</sup> to generate all statistics of interest (such as variance ratios and correlations). We report mean estimates and their corresponding 95% HDI<sup>55</sup>. Code for these analyses are provided on the Open Science Framework at ([osf.io/jsb52](https://osf.io/jsb52)), and we expand on their descriptions in the supplementary materials.

**Polarization and extremism.** We first compared the distributions of information collected by participants against the distributions we would expect from a random, representative sample across trials. These distributions can be seen in Fig. 5. The blue histograms represent the distributions of information sampled by participants, while the orange line represents the true (standard normal) distribution from which the samples were drawn. The difference is immediately striking—as in the re-analyzed studies and the model prediction



**Figure 5.** Distribution of information collected by participants (blue histogram) in the choice (left) and estimation (right) conditions, compared against the expected distribution that would result from unbiased, representative sampling (orange line).

shown in Fig. 1, there is a “dip” in the middle of the distribution of information collected by participants in the choice condition. This bimodal distribution of information shows that people are under-representing moderate information when sampling.

By comparison, the estimation condition shows little to no substantial deviation from the expected distribution, indicating that the estimation intervention appears to evoke more representative sampling than the choice condition. That is, moderate information was considered equally in making estimations, resulting in a typical bell-shaped distribution of information that reflected the true underlying information in the environment.

Our quantitative measure of polarization (KLD) also aligned with these visual observations: the expected KLD in the decision condition was approximately 0.12, but the observed KLD was 1.00 ( $M_{diff} = 0.88$ , 95% HDI = [0.80, 0.95]). The expected KLD in the estimation condition was approximately 0.12, whereas the observed KLD was 0.52 ( $M_{diff} = 0.40$ , 95% HDI = [0.32, 0.47]). Thus, the deviation from the true distribution—and thus the amount of polarization, as defined by KLD—was more than twice as large in the decision condition compared to the estimation condition. KLD scores in the estimation condition were on average .26 standard deviations lower (95% HDI = [−0.30, −0.23]) than KLD scores in the choice condition, corresponding to a medium-sized effect<sup>64</sup>. A model comparison revealed that the regression yielded extremely strong support for inclusion of condition as a factor predicting KLD ( $\log(BF) = 358.30$ ) (Note that we report natural log Bayes factors here and elsewhere in the text, as often the Bayes factors on raw odds scales are approach infinite values. A Bayes factor of 30 on a raw scale (+3.40 on a natural log scale) indicates extremely strong support for an effect/inclusion of a model parameter, while a Bayes factor of 1/30 on a raw scale (−3.40 on a natural log scale) indicates strong evidence against an effect/inclusion of a model parameter.<sup>65,66</sup>

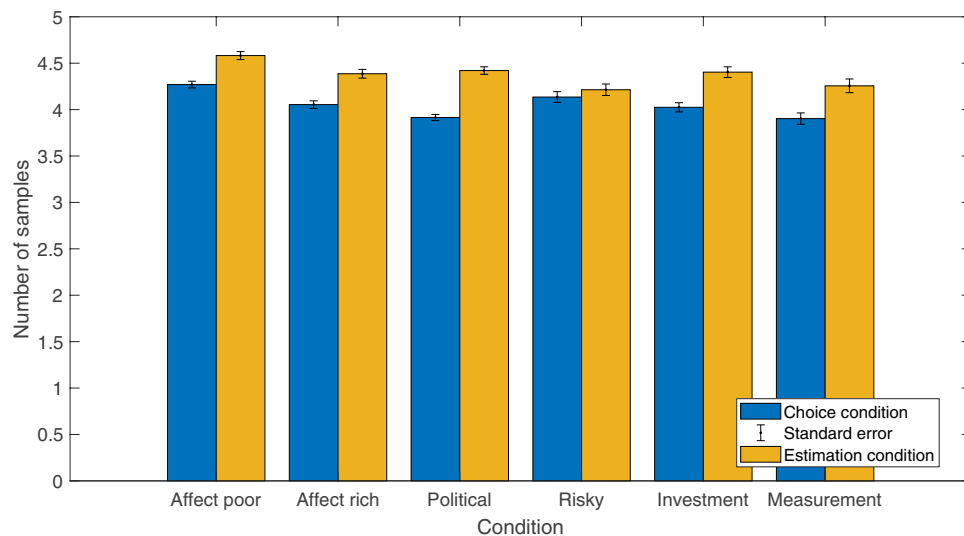
Turning to extremism metrics, we found the condition-level variance ratio for the choice condition to be  $M = 1.18$  (95% HDI = [1.11, 1.27]), where a value of 1.00 indicates that the observed sample is equally as extreme as a representative sample. Conversely, in the estimation condition, the variance ratio was centered right around the expected variance,  $M = 1.01$  (95% HDI = [0.96, 1.09]). These HDIs are entirely non-overlapping, meaning that we can be very confident that the variance in the choice condition is higher than the variance in the estimation condition, signaling that the estimation task reduced the extremeness of information participants gathered.

We also evaluated patterns of extremeness by calculating correlations between the variance of samples people collected with the total number of times they chose to sample information (i.e., number of clicks). The extremeness of information participants collected was credibly correlated with the number of times people sampled information in each trial, both in the choice condition ( $r = -0.44$ ; HDI = [−0.47, −0.41];  $\ln(BF) = 296.16$ ), as well as in the estimation condition ( $r = -0.48$ ; HDI = [−0.51, −0.46];  $\ln(BF) = 397.94$ ). That is to say, the more people sampled extreme information, the fewer times they clicked to receive new information *within both conditions*. The amount (or lack of) of information collected is therefore predictive of extremism above and beyond choice-based polarization, suggesting that manipulations increasing information sampling are likely to be effective in both choice and estimation tasks.

**Information search across conditions.** We hypothesized that one of the reasons for extremism and polarization is that some participants tend to gather less information than others (lower thresholds), and less information typically leads to more volatile/extreme samples of evidence. It is possible that the choice/estimation manipulation, or the question types, led participants to seek out more or less information. This can be examined by looking at the number of pieces of information that participants collected in each condition, quantified as the number of times they clicked on the button to sample more information before making their choice or estimate. A comparison between the different question types, as well as between choice and estimation conditions, on the number of pieces of information sampled is shown in Fig. 6.

To formally quantify whether sampling differed between choice and estimation conditions, we examined the number of samples that participants gathered in each condition. We ran a Bayesian Linear Regression in





**Figure 6.** Mean number of samples drawn (error bars indicate standard error) for each question type in the choice (blue) and estimation (yellow) condition.

Term	Estimate	SE	95% HDI
<b>Information sampling</b>			
Intercept	4.14	0.23	[3.70, 4.60]
Dogmatism	-0.01	0.00	[-0.01, -0.00]
PNS-DFS	-0.14	0.01	[-0.16, -0.12]
PNS-RLS	0.23	0.01	[0.20, 0.26]
<b>Polarization</b>			
Intercept	47.22	2.90	[41.43, 52.99]
Dogmatism	-0.09	0.05	[-0.18, -0.00]
PNS-DFS	0.55	0.26	[0.04, 1.08]
PNS-RLS	-0.70	0.36	[-1.42, -0.01]

**Table 2.** Estimates of the effects of Dogmatism, Personal Need for Structure Reaction to Lack of Structure subscale (PNS-RLS), and Personal Need for Structure Desire for Structure subscale (PNS-DFS) on polarization (K-L divergence, bottom) and information search (number of clicks on the sampling button on each trial, top).

JASP<sup>62</sup> with default/standard priors<sup>67,68</sup>, which provided strong evidence that condition should be included as a predictor of the number of clicks ( $\ln(BF) = 13.12$ ). In the estimation condition, number of clicks increased on average by 0.32 clicks relative to the choice condition (95% HDI = [0.21, 0.44]). Therefore, part of the way that the estimation condition reduces extremism is by encouraging participants to seek out more information, resulting in more informed judgments.

Similarly, a greater number of clicks led to less polarization, reducing the KLD by an average of 0.06 standard deviations per click in the choice condition (95% HDI = [-0.07, -0.05];  $\ln(BF) = 125.29$ ) and by an average of 0.09 standard deviations per click in the estimation condition (95% HDI = [-0.09, -0.08],  $\ln(BF) = 284.67$ ). Gathering more information is a clear-cut strategy for reducing both polarization and extremism, which is why the increase in information search in the estimation condition was so effective at reducing both effects.

To examine whether the type of inference domain affected sampling behavior, we tested whether condition-level manipulations impacted information sampling by looking at the effect of question type (affect poor, affect rich, etc) on number of clicks. We ran a Bayesian ANOVA in JASP<sup>62,69</sup> which indicated that question type had no effect on number of clicks ( $BF(H_0 > H_1) = 270.66$ , indicating strong support for the null hypothesis). There were also no substantial differences on question type in predicting individual-level KLDs ( $BF(H_0 > H_1) = 1237.94$ , indicating strong support for the null hypothesis) so we do not examine question type further.

**Individual differences.** In addition to the differences between conditions, we hypothesized that information search and polarization/extremism should be predicted by personality factors like dogmatism (DOG) and personal need for structure (PNS), which is comprised of the desire for structure (DFS) and reaction to lack of structure (RLS). As summarized in Table 2, the model predicting information search indicated a negative effect of dogmatism, positive effect of the personal need for structure—desire for structure subscale, and negative

effect of the reaction to lack of structure subscale. This aligns with recent findings showing that dogmatism is related to decreased information search under uncertainty<sup>27,59</sup>. Interestingly, the sub-scales of Personal Need for Structure predicted information search in opposite directions. It appears that a desire for structure leads participants to search for less information, perhaps because making a quick decision satisfies the desire to resolve uncertainty and reach a conclusion, while the discomfort associated with having ill-informed views based on little evidence (reaction to lack of structure) drives people to sample more.

These individual differences also predicted individual-level polarization. This was tested by again using a Bayesian linear regression to assess the credibility of individual differences as predictors of participant-level (standardized) KLD. As with information search, the regression suggested that all three measures predicted KLD: dogmatism led to decreased polarization, desire for structure led to greater polarization, and reaction to lack of structure led to less polarization.

The KLD results are particularly noteworthy because dogmatism predicts behavior in the opposite direction than we expected: Greater dogmatism predicted less polarization in resulting sampling behavior (lower KLD), while the PNS scales align with the information search results. So while dogmatism results in reduced information search, this reduction in information sampling does not translate directly to less representative views. One possibility is that highly dogmatic individuals hit a “sampling floor” and sampled only one piece of information before deciding. This would result in a low KLD, because lone samples will not favor extreme information more than moderate information and will, in fact, constitute representative simple random samples. By setting very low thresholds, dogmatic individuals therefore shift back toward a representative sample. As a result, they are ill-informed but no more polarized than a random sample of information from the environment would expect.

## Discussion

Our studies illustrate that regardless of the domain—spanning perceptual choices about dots or colors, preferential choices about food or posters (as revealed in the reanalysis of seven existing studies), neutral social policies, or politically and affectively charged decisions (as revealed in the empirical studies reported here)—the process of decision making generates polarization and extremism during the information sampling process. These were scenarios where there was no social interaction among participants, and where there should be little motivation to engage in biased information processing. While a tremendous volume of work has gone into understanding motivated reasoning<sup>70</sup>, communication dynamics and social networks<sup>21,28,29</sup>, and the social forces that shape and exacerbate polarization<sup>31,32,71,72</sup>, it has focused less on individual-level information search mechanisms that might create polarization before beliefs impacted by social influence. Our work suggests that the root causes of polarization may be even deeper than previously thought, arising from completely rational, individual-level decision strategies.

As we show, polarized and extreme individuals may be those who are merely reacting rationally to choice incentives, and are not necessarily intentionally malicious actors. Based on the model simulations, we should expect that extremists will show up most frequently when there is time pressure to make choices and when decision makers begin the choice process with a biased position. Nevertheless, rational decision making processes—resulting from dual pressures on decision quality and time—will create polarized groups of decision makers and generate a subset of uninformed individuals who hold extreme views.

A major concern regarding these extremists is the degree of influence that they can carry in cultural discourse and the formation of public opinion. Decision makers with small, extreme samples of information are still confident in their views<sup>73,74</sup>, possibly due to a belief in the “law of small numbers” where people think even very small samples will possess the same statistical properties as large samples<sup>75</sup>. This suggests that extremists will have strong convictions and be willing to spread their views, despite the information they have gathered possessing relatively poor reliability. This is exacerbated by several other established phenomena. First, those with the most extreme positions carry a disproportionately greater sway over group discourse because their outlying beliefs are more robust to change and can thus “pull” other people toward them<sup>76</sup>. Second, extremists will be the first to share the information they have—by virtue of taking less time to make decisions, they are free to share their views before those who have taken more time and carefully constructed their beliefs with large samples of information. Extremists will therefore be the first to influence their peers via word of mouth and social networks<sup>77</sup>, provide reviews of new products that influence many subsequent buyer decisions<sup>78</sup>, and spread hateful views that carry quickly through social networks<sup>79</sup>. Social media algorithms base recommendations on early posters, meaning that their “hot takes” will carry greater weight in determining subsequent users’ views and activity as well<sup>77,80</sup>.

As a result, we arrive at a scenario where the least-informed (fewest samples of information), most dogmatic, and most impatient (lowest-threshold) individuals in a population will have the greatest influence on public opinion. The high profile of extremists can explain why perceived political polarization is exaggerated—people tend to think that the political left and right in the US are far more polarized than they are<sup>81</sup>.

**Alleviating polarization and extremism.** The estimation incentives we examined in the new study suggest that polarization and extremism can be reduced at the outset by changing the optimal sampling strategy. Simply asking people to give their best estimate of the difference between options, as opposed to deciding between them, drives a more natural information sampling approach that reduces polarization and extremism while promoting greater information search. Successful implementation of this type of intervention provides avenues for reducing the divides between groups of people, and potentially can be brought to bear in designing simple interventions or judgment architectures to reduce these social problems<sup>82</sup>.

In the political sphere, our results suggest that binary or multiple-choice voting—requiring voters to decide in favor of only Candidate A or Candidate B (or C, D, E, etc)—may be partly responsible for the current state of political polarization. Fortunately, the estimation results suggest a remedy for this particular component of

polarization. Specifically, it may be possible to alleviate political polarization through approaches like cardinal voting<sup>83</sup>, which incentivizes precise ratings (estimation) as opposed to forcing voters to select one option over another. Based on our estimation results, there is reason to hope that this could encourage greater information search among voters as well as more representative sampling. Certainly future work ought to evaluate the practicality and efficacy of different voting interventions like these that could reduce polarization on social issues.

An intriguing possibility that is raised by this work is that many types of information search may yield polarized samples. This may apply beyond literal information search—as we presented here—into domains like information search from memory, which is thought to leverage analogous mechanisms to external information search<sup>38,84</sup>. If search through memory is terminated using the same rules as diffusion or random walk models, then a person's judgment goals at the time of measurement (when they are asked to make a choice or estimate) may determine what information they consider. Similarly, it may be possible to override samples of biased information that have been collected by changing the communication process. Asking a polarized individual to share information to someone who's goal it is to *estimate* may incentivize sharing of representative information from memory. It may therefore be possible to alleviate polarization and extremism by manipulating inference and communication goals in more applied situations like the ballot box, boardroom, or social media feed. To the extent that the interventions we design here are useful for manipulating information search in these domains, they are promising for reducing polarization and extremism.

**Conclusions.** In a decision scenario that incentivizes a trade-off between time and decision quality, rational decision makers will sample more polarized and extreme information than their environment naturally provides. However, we have also identified a potential solution that is as simple as changing decision goals: Polarization and extremism disappear when people are asked to estimate some relative quantity or preferability instead of choosing between options. This approach to making inferences encourages information search while simultaneously reducing the biases inherent to decision-making. Mathematical models, simulations, re-analysis of data from seven empirical studies, as well as results from a new experiment point to the advantages of estimation tasks over choice/decision tasks if one wants to reduce polarization and extremism. Future work will determine how widely and effectively this type of intervention can be implemented to reduce polarization and extremism outside the laboratory, but we hope that this work sheds light on one route to responsibly construct inference environments to create more common ground that brings people together instead of driving them apart.

Received: 9 November 2021; Accepted: 19 April 2022

Published online: 05 May 2022

## References

- Baldassarri, D. & Gelman, A. Partisans without constraint: Political polarization and trends in American public opinion. *Am. J. Sociol.* **114**, 408–446 (2008).
- Wilson, A. E., Parker, V. & Feinberg, M. Polarization in the contemporary political and media landscape. *Curr. Opin. Behav. Sci.* **34**, 223–228 (2020).
- Berger, J. M. *Extremism* (MIT Press, 2018).
- Sunstein, C. R. The law of group polarization. *University of Chicago Law School, John M. Olin Law & Economics Working Paper* (1999).
- Shi, Y., Mast, K., Weber, I., Kellum, A. & Macy, M. Cultural fault lines and political polarization. In *Proceedings of the 2017 ACM on Web Science Conference*, 213–217 (2017).
- Baron, R. S. & Roper, G. Reaffirmation of social comparison views of choice shifts: Averaging and extremity effects in an autokinetic situation. *J. Pers. Soc. Psychol.* **33**, 521 (1976).
- Isenberg, D. J. Group polarization: A critical review and meta-analysis. *J. Pers. Soc. Psychol.* **50**, 1141 (1986).
- Festinger, L. *A Theory of Cognitive Dissonance* Vol. 2 (Stanford University Press, 1957).
- Nickerson, R. S. Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* **2**, 175–220 (1998).
- Westerwick, A., Johnson, B. K. & Knobloch-Westerwick, S. Confirmation biases in selective exposure to political online information: Source bias vs. content bias. *Commun. Monogr.* **84**, 343–364 (2017).
- Lord, C. G., Ross, L. & Lepper, M. R. Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *J. Pers. Soc. Psychol.* **37**, 2098 (1979).
- Iyengar, S., Sood, G. & Lelkes, Y. Affect, not ideology: A social identity perspective on polarization. *Public Opin. Q.* **76**, 405–431 (2012).
- Kuhn, D. & Lao, J. Effects of evidence on attitudes: Is polarization the norm?. *Psychol. Sci.* **7**, 115–120 (1996).
- Gershman, S. J. How to never be wrong. *Psychon. Bull. Rev.* **26**, 13–28 (2019).
- Taber, C. S., Cann, D. & Kucsova, S. The motivated processing of political arguments. *Polit. Behav.* **31**, 137–155 (2009).
- Cook, J. & Lewandowsky, S. Rational irrationality: Modeling climate change belief polarization using Bayesian networks. *Topics Cogn. Sci.* **8**, 160–179 (2016).
- Dixit, A. K. & Weibull, J. W. Political polarization. *Proc. Natl. Acad. Sci.* **104**, 7351–7356 (2007).
- Botvinik-Nezer, R., Jones, M. & Wager, T. Fraud beliefs following the 2020 us presidential election: A belief systems analysis. *PsyArXiv* (2021).
- Wheeler, N. E. *et al.* Ideology and predictive processing: Coordination, bias, and polarization in socially constrained error minimization. *Curr. Opin. Behav. Sci.* **34**, 192–198 (2020).
- Zmigrod, L. Mental computations of ideological choice and conviction: The utility of integrating psycho-economics and bayesian models of belief. *Psychol. Inquiry* (in press).
- Kashima, Y., Perfors, A., Ferdinand, V. & Pattenden, E. Ideology, communication and polarization. *Philos. Trans. R. Soc. B* **376**, 20200133 (2021).
- Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A. & Van Bavel, J. J. Emotion shapes the diffusion of moralized content in social networks. *Proc. Natl. Acad. Sci.* **114**, 7313–7318 (2017).
- Crockett, M. J. Moral outrage in the digital age. *Nat. Hum. Behav.* **1**, 769–771 (2017).

24. Bail, C. A. *et al.* Exposure to opposing views on social media can increase political polarization. *Proc. Natl. Acad. Sci.* **115**, 9216–9221 (2018).
25. Coenen, A. & Gureckis, T. The distorting effect of deciding to stop sampling. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society* (eds Papafragou, A. *et al.*) 2819–2824 (Cognitive Science Society, 2016).
26. Coenen, A. & Gureckis, T. The distorting effects of deciding to stop sampling information. *PsyArXiv* (2021).
27. Zmigrod, L., Eisenberg, I. W., Bissett, P. G., Robbins, T. W. & Poldrack, R. A. The cognitive and perceptual correlates of ideological attitudes: A data-driven approach. *Philos. Trans. R. Soc. B* **376**, 20200424 (2021).
28. Burnstein, E. & Vinokur, A. Testing two classes of theories about group induced shifts in individual choice. *J. Exp. Soc. Psychol.* **9**, 123–137 (1973).
29. Burnstein, E. & Vinokur, A. Persuasive argumentation and social comparison as determinants of attitude polarization. *J. Exp. Soc. Psychol.* **13**, 315–332 (1977).
30. DeKay, M. L. Predecisional information distortion and the self-fulfilling prophecy of early preferences in choice. *Curr. Dir. Psychol. Sci.* **24**, 405–411 (2015).
31. Hogg, M. A. & Turner, J. C. Interpersonal attraction, social identification and psychological group formation. *Eur. J. Soc. Psychol.* **15**, 51–66 (1985).
32. Suhay, E. Explaining group influence: The role of identity and emotion in political conformity and polarization. *Polit. Behav.* **37**, 221–251. <https://doi.org/10.1007/s11109-014-9269-1> (2015).
33. Benkler, Y., Faris, R. & Roberts, H. *Network propaganda: Manipulation, disinformation, and radicalization in American politics* (Oxford University Press, 2018).
34. Olsson, E. J. A Bayesian simulation model of group deliberation and polarization. In *Bayesian Argumentation*, 113–133 (Springer, 2013).
35. Jern, A., Chang, K.-M.K. & Kemp, C. Belief polarization is not always irrational. *Psychol. Rev.* **121**, 206 (2014).
36. Pallavicini, J., Hallsson, B. & Kappel, K. Polarization in groups of Bayesian agents. *Synthese* **198**, 1–55 (2021).
37. Singer, D. J. *et al.* Rational social and political polarization. *Philos. Stud.* **176**, 2243–2267 (2019).
38. Ratcliff, R., Smith, P. L., Brown, S. D. & McKoon, G. Diffusion decision model: Current issues and history. *Trends Cogn. Sci.* **20**, 260–281 (2016).
39. Busemeyer, J. R., Gluth, S., Rieskamp, J. & Turner, B. M. Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends Cogn. Sci.* **23**, 251–263 (2019).
40. Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and humans can optimally accumulate evidence for decision-making. *Science* **340**, 95–98 (2013).
41. Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* **113**, 700–765. <https://doi.org/10.1037/0033-295X.113.4.700> (2006).
42. Edwards, W. Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *J. Math. Psychol.* **2**, 312–329. [https://doi.org/10.1016/0022-2496\(65\)90007-6](https://doi.org/10.1016/0022-2496(65)90007-6) (1965).
43. Wald, A. & Wolfowitz, J. Bayes solutions of sequential decision problems. *Proc. Natl. Acad. Sci. USA* **35**, 99–102 (1949).
44. Tajima, S., Drugowitsch, J. & Pouget, A. Optimal policy for value-based decision-making. *Nat. Commun.* **7**, 1–12 (2016).
45. Boehm, U., van Maanen, L., Evans, N. J., Brown, S. D. & Wagenmakers, E.-J. A theoretical analysis of the reward rate optimality of collapsing decision criteria. *Atten. Percept. Psychophys.* **82**, 1520–1534 (2020).
46. Heitz, R. P. & Schall, J. D. Neural mechanisms of speed-accuracy tradeoff. *Neuron* **76**, 616–628. <https://doi.org/10.1016/j.neuron.2012.08.030> (2012).
47. Wickelgren, W. A. Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol.* **41**, 67–85. [https://doi.org/10.1016/0001-6918\(77\)90012-9](https://doi.org/10.1016/0001-6918(77)90012-9) (1977).
48. Alves, H., Koch, A. & Unkelbach, C. A cognitive-ecological explanation of intergroup biases. *Psychol. Sci.* **29**, 1126–1133 (2018).
49. Rottman, J. & Kelemen, D. Aliens behaving badly: Children’s acquisition of novel purity-based morals. *Cognition* **124**, 356–360 (2012).
50. Gupta, P. *et al.* Space aliens and nonwords: Stimuli for investigating the learning of novel word-meaning pairs. *Behav. Res. Methods Instrum. Comput.* **36**, 599–603 (2004).
51. Decker, J. H., Otto, A. R., Daw, N. D. & Hartley, C. A. From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychol. Sci.* **27**, 848–858 (2016).
52. Zajonc, R. B. & Markus, H. Affective and cognitive factors in preferences. *J. Consumer Res.* **9**, 123–131 (1982).
53. Pachur, T., Hertwig, R. & Wolkewitz, R. The affect gap in risky choice: Affect-rich outcomes attenuate attention to probability information. *Decision* **1**, 64 (2014).
54. Wagenmakers, E. J., Lodewyckx, T., Kuriyal, H. & Grasman, R. Bayesian hypothesis testing for psychologists: A tutorial on the Savage-Dickey method. *Cogn. Psychol.* **60**, 158–189. <https://doi.org/10.1016/j.cogpsych.2009.12.001> (2010).
55. Kruschke, J. K. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and STAN* (Academic Press, 2014).
56. Bowman, A. W. & Azzalini, A. *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations* Vol. 18 (OUP Oxford, 1997).
57. Altemeyer, B. Dogmatic behavior among students: Testing a new measure of dogmatism. *J. Soc. Psychol.* **142**, 713–721 (2002).
58. Neuberg, S. L. & Newsom, J. T. Personal need for structure: Individual differences in the desire for simpler structure. *J. Pers. Soc. Psychol.* **65**, 113 (1993).
59. Schulz, L., Rollwage, M., Dolan, R. J. & Fleming, S. M. Dogmatism manifests in lowered information search under uncertainty. *Proc. Natl. Acad. Sci.* **117**, 31527–31534 (2020).
60. Evans, N. J., Rae, B., Bushmakina, M., Rubin, M. & Brown, S. D. Need for closure is associated with urgency in perceptual decision-making. *Mem. Cogn.* **45**, 1193–1205 (2017).
61. Team, R. C. R: A language and environment for statistical computing (2013).
62. JASP Team. JASP (Version 0.14.1)[Computer software] (2020).
63. Plummer, M. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, Vol. 124, 10 (2003).
64. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences* (Academic Press, 2013).
65. Jeffreys, H. *The Theory of Probability* (OUP Oxford, 1961).
66. Raftery, A. E. Bayesian model selection in social research. *Sociol. Methodol.* **25**, 111–63 (1995).
67. Clyde, M. A., Ghosh, J. & Littman, M. L. Bayesian adaptive sampling for variable selection and model averaging. *J. Comput. Graph. Stat.* **20**, 80–101 (2011).
68. Consonni, G., Fouskakis, D., Liseo, B. & Ntzoufras, I. Prior distributions for objective Bayesian analysis. *Bayesian Anal.* **13**, 627–679 (2018).
69. Rouder, J. N., Morey, R. D., Speckman, P. L. & Province, J. M. Default Bayes factors for ANOVA designs. *J. Math. Psychol.* **56**, 356–374 (2012).
70. Rigoli, F. Masters of suspicion: A Bayesian decision model of motivated political reasoning. *J. Theory Soc. Behav.* **51**, 350–370 (2021).
71. Tajfel, H. E. *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations* (Academic Press, 1978).

72. Tajfel, H. E. & Turner, J. C. An integrative theory of intergroup conflict. In (M. A. Hogg & D. Abrams Eds.) *Intergroup relations: Essential readings*. 94–109 (2001).
73. Kvam, P. D. & Pleskac, T. J. Strength and weight: The determinants of choice and confidence. *Cognition* **152**, 170–380. <https://doi.org/10.1016/j.cognition.2016.04.008> (2016).
74. Griffin, D. & Tversky, A. The weighing of evidence and the determinants of confidence. *Cogn. Psychol.* **24**, 411–435 (1992).
75. Tversky, A. & Kahneman, D. Belief in the law of small numbers. *Psychol. Bull.* **76**, 105 (1971).
76. Navarro, D. J., Perfors, A., Kary, A., Brown, S. D. & Donkin, C. When extremists win: Cultural transmission via iterated learning when populations are heterogeneous. *Cogn. Sci.* **42**, 2108–2149 (2018).
77. Lynn, T., Muzellec, L., Caemmerer, B. & Turley, D. Social network sites: Early adopters' personality and influence. *J. Prod. Brand Manag.* **26** 42–51 (2017).
78. Stephen, A. T. & Lehmann, D. R. How word-of-mouth transmission encouragement affects consumers' transmission decisions, receiver selection, and diffusion speed. *Int. J. Res. Market.* **33**, 755–766 (2016).
79. Mathew, B., Dutt, R., Goyal, P. & Mukherjee, A. Spread of hate speech in online social media. In *Proceedings of the 10th ACM Conference on Web Science*, 173–182 (2019).
80. Bastide, P. R., Deluca, L. S. & Do, L. M. User recommendations in a social media network (2018). US Patent 9,871,758.
81. Westfall, J., Van Boven, L., Chambers, J. R. & Judd, C. M. Perceiving political polarization in the United States: Party identity strength and attitude extremity exacerbate the perceived partisan divide. *Perspect. Psychol. Sci.* **10**, 145–158 (2015).
82. Thaler, R. H. & Sunstein, C. R. *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Penguin, 2009).
83. Baujard, A., Gavrel, F., Igersheim, H., Laslier, J.-F. & Lebon, I. How voters use grade scales in evaluative voting. *Eur. J. Polit. Econ.* **55**, 14–28 (2018).
84. Shiffrin, R. M. Memory search. *Models of Human Memory* 375–447 (1970).

### Author contributions

A.A., C.E.M., and A.M. created the stimuli and collected data, P.D.K. and A.A. analyzed the data., P.D.K. and M.B. conceived of the study and conditions in the main text, and all authors contributed to writing the paper.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-11389-0>.

**Correspondence** and requests for materials should be addressed to P.D.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022