



# Collective wisdom in polarized groups



Collective Intelligence  
August-September 2022: 1–14  
© The Author(s) 2022  
Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: 10.1177/26339137221104788  
[journals.sagepub.com/home/col](https://journals.sagepub.com/home/col)  
 SAGE

**Joseph B Bak-Coleman**

Center for an Informed Public, University of Washington, Seattle, WA, USA; Science Institute, University of Washington, Seattle, WA, USA

**Christopher K Tokita**

Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA

**Dylan H Morris**

Department of Ecology & Evolutionary Biology, UCLA

**Daniel I Rubenstein**

Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA

**Iain D Couzin**

Department of Collective Behaviour, Max Planck Institute for Animal Behavior, Konstanz, Germany; Center for the Advanced Study of Collective Behaviour, University of Konstanz, Konstanz, Germany; Department of Biology, University of Konstanz, Konstanz, Germany

## Abstract

The potential for groups to outperform the cognitive capabilities of even highly skilled individuals, known as the “wisdom of the crowd”, is crucial to the functioning of democratic institutions. In recent years, increasing polarization has led to concern about its effects on the accuracy of electorates, juries, courts, and congress. While there is empirical evidence of collective wisdom in partisan crowds, a general theory has remained elusive. Central to the challenge is the difficulty of disentangling the effect of limited interaction between opposing groups (homophily) from their tendency to hold opposing viewpoints (partisanship). To overcome this challenge, we develop an agent-based model of collective wisdom parameterized by the experimentally-measured behaviour of participants across the political spectrum. In doing so, we reveal that differences across the political spectrum in how individuals express and respond to knowledge interact with the structure of the network to either promote or undermine wisdom. We verify these findings experimentally and construct a more general theoretical framework. Finally, we provide evidence that incidental, context-specific differences across the political spectrum likely determine the impact of polarization. Overall, our results show that whether polarized groups benefit from collective wisdom is generally predictable but highly context-specific.

## Keywords

collective wisdom, polarization, wisdom of the crowds

## Corresponding author:

Joseph B Bak-Coleman, Center for an Informed Public, University of Washington, Mary Gates Hall, Seattle, WA 98195, USA.

Email: [joebak@uw.edu](mailto:joebak@uw.edu)



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

## Significance Statement

Whether or not individuals in groups benefit from the “Wisdom of the Crowd” is a fundamental sociological question. Unfortunately, our understanding of collective wisdom relies on restrictive assumptions that are a far cry from the homophilous, partisan, and demographically diverse conditions in which many decisions are made. Here, we examine the impact of partisanship and homophily in the context of the US political system. Using experimental and computational approaches, we find that heterogeneity, homophily and partisanship can impact collective wisdom, even in the context of apolitical trivia questions. We validate our model experimentally, and provide evidence that consequences of polarization may vary widely across knowledge domains. Finally, we provide a general mathematical model that frames our findings and highlights the potential for similar effects in other domains.

## Introduction

Early theories of collective wisdom relied on the unrealistic assumption that members of a group act independently (i.e. do not influence one another). Yet, deliberation is a hallmark of democracies, taking place everywhere from dinner tables to congress and more recently, online. Incorporating social interactions into our understanding of when and whether collective wisdom will occur has been an increasing focus of both theoretical and empirical work (Galam 2002; Becker et al. 2019, 2017; O'Connor and Weatherall 2018; Davis-Stober et al., 2014; Jayles et al., 2017).

Intuitively, when a group faces a discrete choice between two options with one being correct, deliberation will be beneficial if it tends to convert incorrect opinions to correct ones. In this case, assuming members are identical makes it possible to study collective wisdom using frameworks of opinion dynamics, social epistemology, and statistical physics (Li et al., 2013; Moussaïd et al., 2013; Lorenz and Neumann 2019). Recent extensions of these approaches have incorporated polarization wherein subgroups interact preferentially within their group (i.e. homophily) (O'Connor and Weatherall 2018).

However, in addition to homophily, subgroups may generate and/or respond to social information in different ways or differ in their prior likelihood of holding a given belief (i.e. partisanship, issue polarization). Incorporating this nuance into a purely theoretical approach is difficult, as it dramatically increases the number of parameters and frustrates model interpretation. Despite these challenges, theoretical insight is necessary to evaluate the generality of empirical work demonstrating collective wisdom in crowds estimating politically-charged quantities or producing high-quality Wikipedia articles Guilbeault et al. (2018); Becker et al. (2019); Shi et al. (2019).

While promising, experiments on politically-charged topics are constrained by existing partisanship and often it is not possible to evaluate counter-factual conditions. For instance, research has previously considered collective wisdom in the context of projecting risks of climate change Becker et al.

(2019). In this context, conservatives tend to under-estimate risks and liberals tend to over-estimate. Without risking selection bias, an experiment cannot create conditions in which a) conservatives tend to under-estimate and liberals over-estimate a value of interest b) or the reverse is true, and c) where there is no systematic bias across the political spectrum. As a result, it is challenging to experimentally disentangle how collective wisdom is impacted by partisanship (i.e. who holds a belief) versus other cognitive differences across the political spectrum. Such cognitive differences could include susceptibility to social influence, persuasiveness, or ability to assess ones knowledge in a given domain.

Using the US political system as a model, we employ a hybrid theoretical and empirical approach to uncover how partisanship, homophily, and cognitive differences all impact collective wisdom. We begin by asking participants across the political spectrum a set of politically-neutral trivia questions. The participants' responses and political leanings provide data that we used to derive and parameterize a statistical, agent-based, model of collective wisdom. We use this model to make testable predictions about collective wisdom as a topic becomes partisan and/or interactions between conservatives and liberals are rare. The general findings of this model are tested and verified via asynchronous online experiments. Next, we present a formal mathematical model that both reproduces our results and can be parameterized to consider polarized collective wisdom in other sociological and ecological contexts. Finally, we provide evidence that incidental domain-specific relationships between confidence and accuracy are more likely to drive patterns of polarized collective wisdom than deep-seated tendencies.

## Experimental design

A diverse, national sample (Experiment 1:  $N = 394$ , Experiment 2:  $N = 979$ ) of people across the political spectrum were recruited and asked a series of trivia questions about state capitals. This provided a non-partisan set of True/False questions that vary in difficulty while representing a similar form of knowledge. All experiments began with obtaining

self-reported political leanings and demographic data (Figures 5, and 6). Previous work has shown that self-reported political ideology is a reliable predictor of underlying ideological beliefs (Preoiuc-Pietro et al., 2017).

Our experiments were designed to allow us to investigate how collective wisdom is influenced by two distinct phenomena often referred to as polarization. The first, homophily, refers to conditions where frequent/preferential contact between similar individuals results in interactions between conservatives and liberals being rare. The second is the tendency of conservatives and liberals to hold opposing viewpoints, which (for clarity) we refer to as partisanship (Narayanan et al., 2018; Preoiuc-Pietro et al., 2017). From the perspective of the model and experiments, homophily ( $h$ ) describes the frequency of interactions between conservatives and liberals (a network-structural property) whereas partisanship describes the distribution of initial answers across the political spectrum (an individual-level property).

In the first experiment (Experiment 1) we characterize, generally, the way in which people across the political spectrum assess their own knowledge and respond to social information. To do this, participants were asked to answer 20 True/False questions consisting of “X is the capital of Y”, where Y was a U.S. state, and X was the true capital half of the time, and the second largest non-capital city the remainder (based on 2010 US census data, See Methods). For each question, they were asked to likewise drag a slider to indicate the probability they believed their answer to be correct from 50% (chance) to 100%. After answering all 20 questions, participants were then provided the same question, with an additional statement that read “We asked  $N$  other respondents what they thought was the correct answer and how confident they were. Overall, they believed there was a  $p\%$  chance the answer was true and a  $(100 - p)\%$  chance the answer was false”.  $N$  was chosen uniformly between 1 and 15, and  $p$  was chosen uniformly between 0 and 100. All experiments and methods were approved by the Princeton University Institutional Review Board.

A challenge in studying the impact of polarization is that synchronous online experiments would require many large groups of players across the breadth of the political spectrum and across various conditions. Further examining the impact of groups holding opposing viewpoints without confounds introduced by using politically-charged knowledge would require filtering from a very large initial pool. Ultimately, these factors render synchronous online experiments both impractical and prohibitively expensive.

Instead, we relied on the data gathered in Experiment 1 to parameterize an agent-based simulation of collective wisdom in a large group of interacting players. This was done using hierarchical Bayesian inference to parameterize a pair of models that approximated our participants’ behavior in 1) assessing their own knowledge (confidence) and 2) their

tendency to change their guess when confronted with conflicting social information. Each model took into account differences across the political spectrum, allowing us to simulate how individuals behave depending on their specified political leaning. We then investigated collective accuracy by arranging these statistical-agent models in structured social networks. From this we made predictions regarding how collective wisdom is impacted by polarization (i.e. homophily and partisanship) across context and for a range of question difficulties.

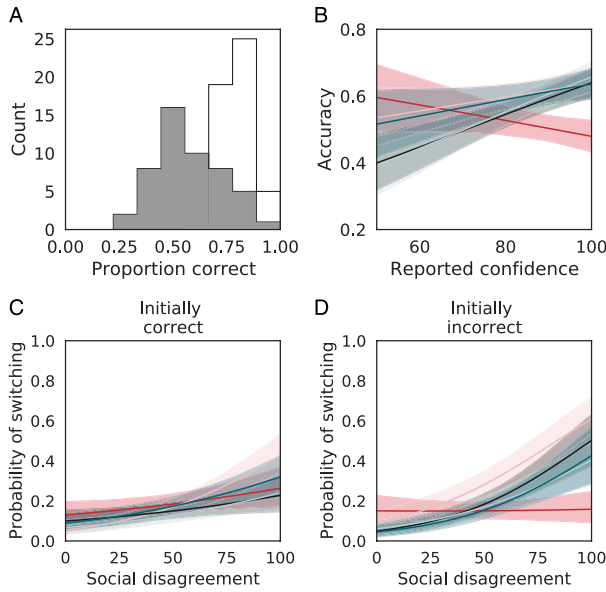
We then tested these predictions using an online experiment (Experiment 2) in which participants were provided social information aggregated from the answers given in the first experiment. Participants in experiment 2 were shown selected social information from participants in the first experiment (See Methods). Selecting social information from past players allowed us to explicitly control homophily,  $h$ , such that participants connected with someone from the same side of the aisle with probability  $h$ , and connected with someone across the aisle with probability  $1 - h$ . It further allowed us to pre-determine the proportion of correct participants on each side of the political spectrum (partisanship) in a way that would not be possible synchronously. For instance, we could construct networks in which the vast majority of conservatives were correct, with liberals being incorrect or vice-versa.

We note that participants were not aware of whether they were correct nor the partisanship of their neighbors. We used this design to examine high and low homophily ( $h = .98, h = .5$ ), as well as extremes of partisanship. We used participants’ answers to estimate accuracy in each condition and evaluate the predictions of the simulation study. In a final experiment (Experiment 3) we evaluated whether observed differences in individual behavior across the political spectrum generalized to a wider range of knowledge domains. The overall presentation was similar to the non-social portion of experiments 1 and 2, yet involved a range of trivia questions unrelated to state capitals (See Methods).

## Results

### Experiment 1

We began by characterizing the relationship between confidence and accuracy across the political spectrum (Figure 1). This can be a key determinant of collective wisdom; within a wise crowd, accurate individuals should express more confidence in their answers than those that are incorrect (Laan et al., 2017; De Polavieja and Madirolas 2014; Almaatouq et al., 2020). Consistent with previous work, “True” questions were answered correctly the vast majority of the time, making it unlikely that collective properties would (or can) alter the likelihood of a majority correct (i.e. wise) response (Condorcet, 1785; Prelec et al.,



**Figure 1.** (A) State-level average accuracy for False (grey) and True (white) questions. (B) Probability being correct as a function of reported confidence. (C) Probability of a participant changing their guess as a function of the social evidence in favor of the opposing answer when initially correct. (D) As in C, when initially incorrect. Figures C–D are posterior estimates from a Bayesian logistic regression (See Methods). Very Conservative: Dark Red, Conservative: Pink, Moderate: Grey, Liberal: Light Blue, Very Liberal: Dark Blue.

2017) (Figure 1A). We accordingly based our analysis on less-intuitive, “False” questions.

We found positive relationships between reported confidence and accuracy for moderate (M) and liberal (L) participants, yet no relationships were observed for conservative (C) and very liberal (VL) participants (Figure 1, Table 3). Very conservative (VC) participants, however, exhibited a negative relationship between accuracy and confidence and were less likely to answer questions correctly (Table 4).

It is important to note that the relationships between confidence and knowledge observed here should not be interpreted as intrinsic to, or caused by, political ideology. We demonstrate that this is not likely to be the case in Experiment 3. Yet, as we show below, even incidental associations between ideology, correctness, and confidence can impact collective wisdom. Further, these findings echo previous research indicating that poor performers tend to over-rate their abilities (Kruger and Dunning 1999), as well as work demonstrating overconfidence among those holding more extreme political beliefs (Ortoleva and Snowberg 2015b,a).

Whether or not a group of individuals is collectively wise depends not only on individual estimates but, should they interact, also how those individuals integrate personal and social information. We used a hierarchical Bayesian logistic

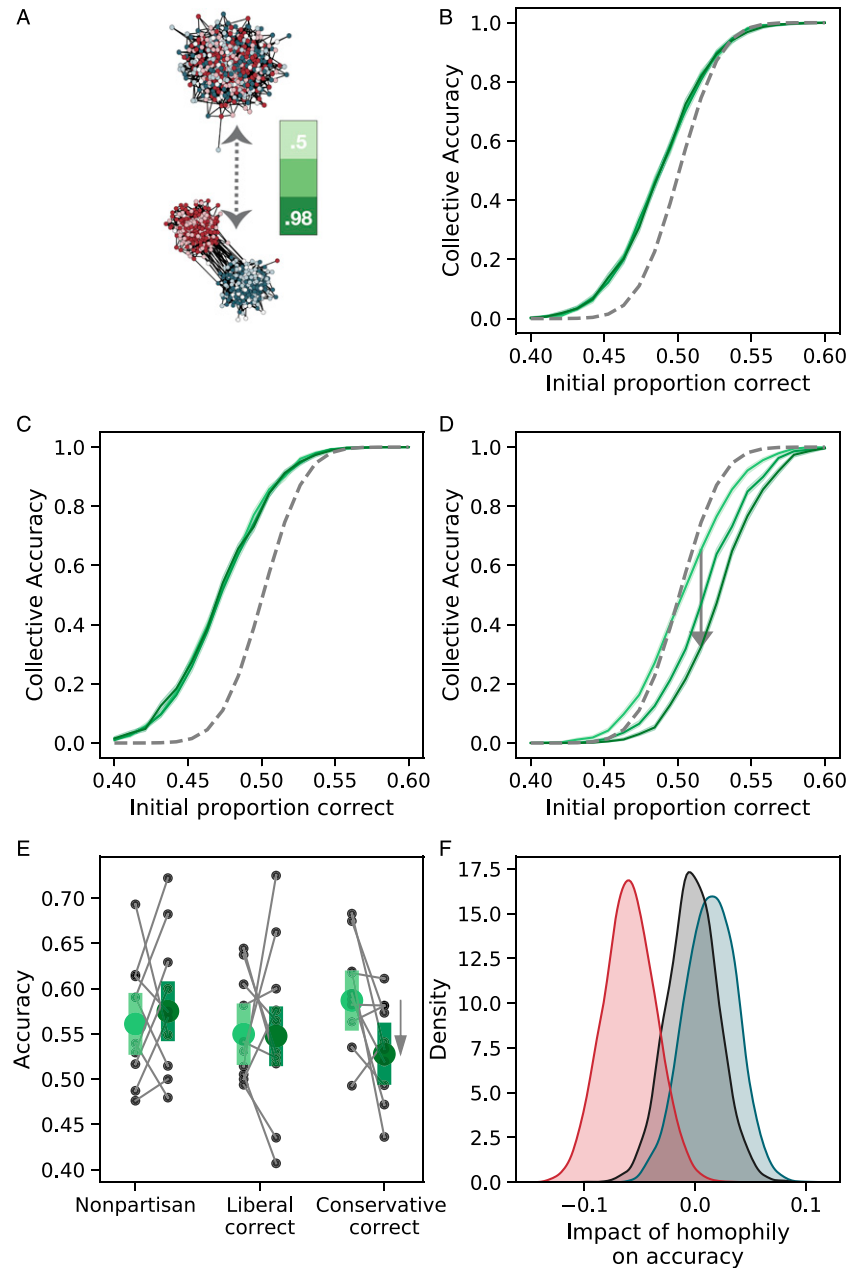
regression to quantify the extent to which social information and confidence impact the likelihood of a participant altering their guess (Figures 1(C) and (D)). We found that social information contrary to a participant’s belief generally encourages them to change their guess, and this effect is reduced proportionately to the confidence they expressed in that initial belief (Table 6 and 7).

Very conservative participants with incorrect initial guesses, however, were not generally influenced by social information ( $\beta_{SB,VC}$ :  $-0.15$ , 89% C.R.:  $[-1.11, 0.82]$ ) (Figure 1, Table 7). As with overconfidence, our analysis only identifies this as an incidental relationship between ideology and responsiveness in a given context. However, as we will show below, questions regarding state capitals provide a reasonable and less politically charged set of parameters that can reveal more general phenomena.

### Simulation study

In order to understand how differences in individual confidence and responsiveness to social information can impact patterns of collective wisdom, and to generate experimentally-testable predictions regarding the effect of polarization on collective wisdom, we constructed an agent-based model parameterized from our survey data. For simplicity, we consider collective wisdom as the probability of a majority coming to the correct decision Condorcet (1785). Briefly, agents were placed in social networks, where they made initial guesses and then received social information from neighbors to whom they were connected. Agents’ confidence in their initial guesses was drawn from a data-parameterized zero-one-inflated beta distribution (SI Figure 9). Their propensity to change their guess when faced with conflicting social information was simulated based on a multi-level logistic regression fit to our human participant survey data (Figures 8, 9, and 10). In our simulations, agents with a given political leaning and initial guess, agents’ reported confidence and response to social information were statistically similar to human study participants. Details and links to all data and code are provided in the methods.

We assigned political leanings to 500 agents based on the relative frequencies observed in online retweet networks (Preoiuc-Pietro et al., 2017) (Figure 2A). Social connections (edges) were added to the network such that conservatives and liberals made connections with like-minded agents with probability  $h$  and connections with agents across the political divide with probability  $1 - h$ . By varying  $h$ , we could look at the impact of moving from a randomly connected network (e.g.  $h = 0.5$ ) to a homophilous one in which conservatives and liberals rarely interact ( $h \approx 1.0$ ). Collective wisdom was simulated by assigning a random subset of the network the correct answer initially (from 40–60%) and examining the resultant wisdom after social interaction and revision of guesses. As the agents’ initial answer (True



**Figure 2.** (A) Visual depiction of heterogeneous ( $h = .5$ ) vs. homophilous ( $h = .98$ ) social networks. (B) No impact of homophily for nonpartisan questions across levels of homophily. From light green to dark green:  $h = .5, .75, .98$ . Shaded regions indicate 89% compatibility intervals (C.I.) (C) Little to no benefit of homophily when liberals tend to initially have the correct answer. (D) When the conservative group tends to be initially correct, homophily leads to decreased collective wisdom (Grey Arrow). (E) Empirically observed average participant accuracy in heterogeneous (light green) and homophilous (dark green) conditions. The grey arrow indicates the predicted and observed decline in wisdom for conservative-correct initial conditions. Dots indicate average accuracy for a given question and shaded lines indicate 89% compatibility intervals for the average across states. (F) Posterior distributions for the effect of homophily on accuracy with increasing homophily.

or False) is determined by the starting conditions, this approach controls for observed differences across the political spectrum in the probability of a participant being correct. We simulated partisan knowledge by setting the

starting conditions of the model such that the majority of conservatives answer “False”, and liberals answer “True” (conservative-correct), vice-versa (liberal-correct) or distribute answers randomly (nonpartisan, described above).



In the simple case where correct answers were equally likely across the political spectrum (i.e. non-partisan) homophily had no discernible impact on collective wisdom (Figure 2(B)). However, under partisan conditions, we found that effects of homophily on collective wisdom depended upon which side of the political spectrum tended to answer correctly. When the liberal portion of the network was correct, homophily had little to no effect on wisdom. Yet, when conservatives had the correct answer homophily lead to a marked decrease in collective accuracy (Figures 2(C) and (D)). In the most extreme cases, networks with high homophily, and where conservatives were correct, were  $\approx 40\%$  less likely to have a majority correct outcome.

## Experiment 2

By using trivia questions related to state capitals, our model was able to isolate the impact of partisanship without relying on politically-charged questions. This further allowed questions to be in the same general knowledge domain, minimizing risks of domain-specific effects. A limitation of this approach, however, is that testing the predictions at scale would require filtering a final group from a much larger group of initial participants. Doing this synchronously would not be tractable, nor practical. Thus, in order to test the predictions of the data-driven model, we instead conducted an asynchronous online experiment (Experiment 2,  $n = 976$ ).

Experiment 2, from the participants perspective, resembled experiment 1. However, social information was drawn from the empirical distributions of answers obtained in experiment 1. Social information shown to participants was weighted by condition (i.e., homophily, partisanship) and the focal participant's political leaning to match social information used by agents networks in the computational model. As described above, our model predicts that homophily will impact collective wisdom, but only when knowledge is partisan. We used a hierarchical binomial regression to quantify the average participant accuracy for each question at two levels of homophily ( $h = .5, h = .98$ ) and in liberal-correct, conservative-correct and nonpartisan networks.

Consistent with our theoretical predictions, we observe that homophily has little to no effect in nonpartisan and liberal-correct initial conditions yet degrades collective wisdom when conservatives tend to be correct (Figure 2(E), Table 8). Specifically, we observed a 6% absolute decrease in participant accuracy (89% CI: 2–10%). We note that the smaller effect size here is to be expected, as we are measuring accuracy of individual participants rather than a collective. In a collective context relying on a simple majority, small changes in individual accuracy can be amplified to large changes in the probability of the majority being correct as group size increases (i.e. Condorcet's Jury Theorem) (Condorcet, 1785).

## Conceptual model

The empirical and computational work described above demonstrates that polarization (e.g. homophily and partisanship) can impact collective wisdom in a predictable way. In other knowledge domains, ideology-specific relationships between confidence and accuracy will likely differ altering how polarization impacts collective wisdom. In order to facilitate extension outside of our experimental context, and develop intuitive insight for the computational and empirical results, we developed a minimal theoretical model of belief updating based on social information in a structured population.

We begin by considering a single population of  $n$  individuals of which a proportion  $q$  are initially correct. The overall amount of social influence is represented by factor  $\lambda$ . Correct individuals convince incorrect individuals with a probability  $\alpha$  per interaction, and the reverse occurs at probability  $\beta$  per interaction (equation (1)).

$$\begin{aligned}\Delta q &= \lambda(\alpha q(1 - q) - \beta q(1 - q)) \\ &= \lambda q(1 - q)(\alpha - \beta)\end{aligned}\quad (1)$$

Intuitively, we will not see a change in collective wisdom if individuals do not interact  $\lambda = 0$ , are entirely wrong  $q = 0$  or are entirely correct  $q = 1$ . However, given  $\lambda > 0$  (individuals interact),  $0 \geq \alpha, \beta \geq 0$  ( $\alpha$  and  $\beta$  are valid probabilities) and  $0 < q < 1$  (both correct and incorrect individuals are initially present), crowds will become more wise ( $\Delta q > 0$ ) when incorrect individuals are on average more frequently converted to being correct ( $\alpha > \beta$ ).

We can extend this model formula to consider two subpopulations of equal size that differ in the extent to which they are initially correct ( $q_1$  and  $q_2$ ) and the degree to which they interact with the other subpopulation ( $h$ , homophily, equation (2)). Table 2 Individuals may also differ in their tendency to convince members within their group and outside of it:  $\alpha_{ij}$  is the per-interaction probability that a correct individual from population  $i$  converts an incorrect individual from population  $j$ , and  $\beta_{ij}$  is the per-interaction probability that an incorrect individual from  $i$  converts a correct member of  $j$ .

$$\begin{aligned}\Delta q &= \frac{\Delta q_1 + \Delta q_2}{2} \\ &= \frac{\lambda h}{2} (q_1(1 - q_1)(\alpha_{11} - \beta_{11}) + q_2(1 - q_2)(\alpha_{22} - \beta_{22})) \\ &\quad + \frac{\lambda(1 - h)}{2} (q_2(1 - q_1)(\alpha_{21} - \beta_{12}) \\ &\quad + q_1(1 - q_2)(\alpha_{12} - \beta_{21}))\end{aligned}\quad (2)$$

In principle, each value of  $\alpha_{ij}$  and  $\beta_{ij}$  can be estimated or measured for a given context. Here we examine a

parameterization that captures a key qualitative aspect of our study population, namely that liberal participants are more willing to switch to the correct answer (Figure 1(D)). We note that no meaningful differences were observed in conversion rates to the incorrect answer (Figure 1(C)). We further assume that conversion rates to and from the incorrect answer for conservatives are similar. We therefore wish to choose values for  $\alpha_{ij}$  and  $\beta_{ij}$  such that:

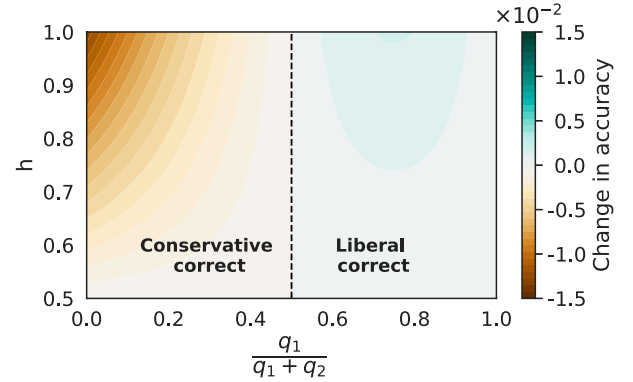
$$\alpha_{1,1} = \alpha_{2,1} > \alpha_{1,2} = \alpha_{2,2} = \beta_{1,1} = \beta_{1,2} = \beta_{2,1} = \beta_{2,2} \quad (3)$$

To satisfy these constraints, we chose values of .25 and .2 for each side of the inequality which is similar to the overall rates of conversion observed in Experiment 1 (Figure 10). This simple parameterization nevertheless captures the key findings from both our simulation study and Experiment 2. For conservative correct initial conditions (e.g.  $\frac{q_1}{q_1+q_2} \approx 0$ ) wisdom monotonically decreases with increasing homophily (Figure 3). Incorrect  $q_1$  individuals are easily swayed by correct  $q_2$  individuals, so collective wisdom declines when increasing homophily prevents such conversion events from taking place. However, when  $\frac{q_1}{q_1+q_2} = 0.5$  (nonpartisan) or  $\frac{q_1}{q_1+q_2} \approx 1$  (liberal-correct) we observe little to no impact of increasing homophily. When  $q_1$  individuals are generally correct, homophily has little effect: correct  $q_1$  individuals cannot effectively convert  $q_2$  individuals, and incorrect  $q_1$  individuals are rare. Detailed analysis can be found in the Supplement. Overall, wisdom is promoted by maximizing contact between persuadable incorrect individuals and correct individuals, while minimizing contact between persuasively incorrect individuals and the rest of the population.

### Experiment 3

The model described above highlights how incidental relationships between confidence, accuracy, partisanship, and homophily come together to alter collective wisdom. Experiments 1 and 2 considered knowledge solely within a single domain, raising the question of whether the observed relationships between confidence and accuracy are general phenomena or specific to a single knowledge domain (i.e. state capitals). To gain insight into this question we conducted an experiment (Experiment 3,  $N=498$ ) in which we asked people across the political spectrum to answer a set of True/False trivia questions requiring knowledge across a wider range of domains (Table 9). As in experiment 1 and 2, we asked participants to report the confidence in addition to each answer.

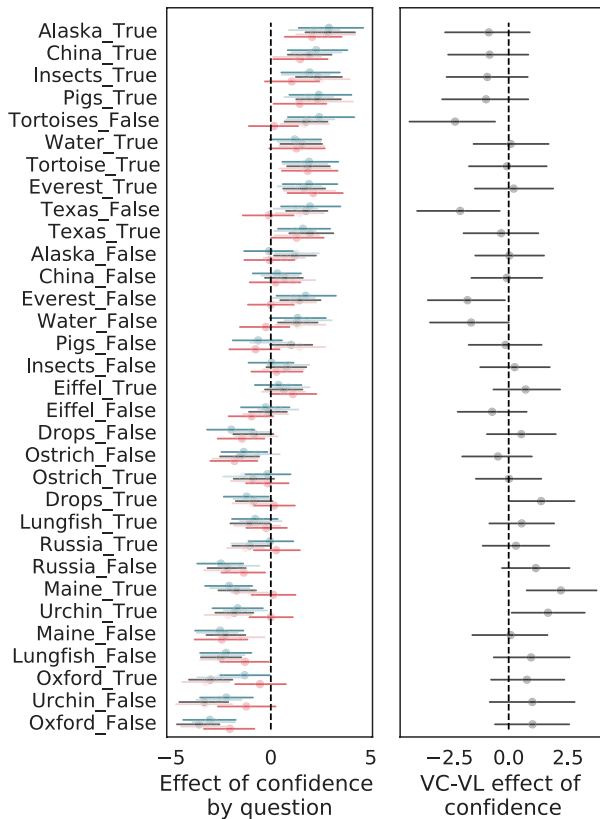
We used a cross-classified hierarchical Bayesian model to examine the extent to which relationships between confidence and accuracy are driven by domain-specific (i.e. question-specific) effects, or properties intrinsic to the individuals and invariant across domains. This analysis



**Figure 3.** Results from our conceptual model. Plotted is the change in average individual accuracy, relative to a well-mixed group as a function of increasing homophily  $h$  and partisanship  $\frac{q_1}{q_1+q_2}$ . Here, the total proportion correct is fixed such that  $\frac{q_1+q_2}{2} = 0.5$   $h$  is the proportion of interactions that occur within vs across groups,  $q_1$ : the proportion of liberals that are initially correct,  $q_2$  the proportion of conservatives that are initially correct. The dashed line represents represents nonpartisan conditions.

revealed that incidental differences across the political spectrum in how confidence relates to accuracy are common (Figure 4, Table 10). However, the direction and magnitude of these differences varied across knowledge domains such that VC participants exhibited more positive relationships than VL participants for some questions, and more negative relationships for others. Despite this question-specific variation, no overall differences in the relationship between confidence and accuracy across the political spectrum were observed ( $\beta_{\text{Conf,VC-VL}}: -0.07$ , 89% C.R.:  $[-0.28, 0.14]$ , Table 11).

Beyond political leaning it remains possible that other individual attributes impact confidence/accuracy relationships. To evaluate this possibility, participants in Experiment 3 were given a series of established cognitive tests to evaluate their Bullshit Receptivity, Verbal intelligence (wordsums), use of Heuristics and Biases, Numeracy, Need for Cognition, and Faith in Intuition (Pennycook et al., 2015; Frederick 2005; Schwartz et al., 1997). Of these, only numeracy ( $\beta_{\text{NUMS}}: 0.09$ , 89% C.R.:  $[0.03, 0.14]$ ) and verbal intelligence ( $\beta_{\text{VI}}: 0.11$ , 89% C.R.:  $[0.05, 0.16]$ , Table 12) predicted individual confidence-accuracy relationships. In both cases, higher scores were associated with more positive relationships between confidence and accuracy. Notably, neither numeracy or verbal intelligence are strongly associated with political ideology (Meisenberg 2015; Choma et al., 2019). Taken together, experiment 3 provides evidence that incidental, domain-specific effects are more likely to shape polarized collective wisdom than a consistent pattern arising from deep-seated differences.



**Figure 4.** Mean and 89% credible intervals for the effect confidence on accuracy for a wide range of trivia questions. Raw effects for each political leaning are plotted on the left, and the difference in the effect between VC and VL participants is plotted on the right. Very Conservative: Dark Red, Conservative: Pink, Moderate: Grey, Liberal: Light Blue, Very Liberal: Dark Blue.

## Discussion

Our results demonstrate that polarization can be a key determinant of collective wisdom in heterogeneous groups. Even in the context of trivia questions about state capitals, we find evidence that the effect of polarization on group correctness can be either negative or neutral. Our computational and mathematical approaches reveal how these results arise; they are produced by the interplay between persuadability, homophily, and partisanship. Notably, deep-seated behavioral differences across the political spectrum are not necessary—incidental, domain-specific relationships between confidence, knowledge, responsiveness, and ideology are sufficient to impact collective wisdom. We further provide evidence that such incidental differences are common, and that deep-seated differences in confidence/accuracy relationships are neither necessary nor likely. We anticipate that these parameters will further vary across political contexts (e.g. online vs offline), knowledge domains, and time. Consequently, polarization should not be

treated as a force that universally improves (or degrades) collective wisdom.

The impact of polarization on decision-making will have context-dependent effects, both in magnitude and in direction. A corollary of this is that experiments on collective wisdom in partisan and polarized contexts are likely to provide apparently conflicting results owing to researcher degrees of freedom (Simmons et al., 2011). For example, pilot experiments with initially broad question sets could be narrowed down to a final experiment that confirms virtually any initial hypothesis. Further, study-averaged effects of polarization on wisdom may be biased composites of particular knowledge domains in which positive, negative, and neutral effects are present.

In light of this, it may be difficult to effectively evaluate the overall effect of political polarization on (un)wise collective decisions. This finding joins a host of others highlighting the ways in which social influence, forced binary choices, complex environments and deliberation can impact the ability of groups to benefit from collective wisdom (Lorenz et al. (2011); Frey and Van de Rijt (2020); Kao and Couzin (2014); Becker et al. (2021, 2020). Nevertheless, given the presence of effects in even neutral knowledge domains, we expect polarization will impact collective decision-making—perhaps more so—in more complex and ideological knowledge domains. In these contexts, identity is likely to play a larger role altering dynamics in unique ways if individuals are aware of their alters’ identity (Guilbeault et al. (2018); Becker et al. (2019).

Extension to these domains may be possible by adjusting the  $\alpha$  and  $\beta$  terms in the theoretical model accordingly. More generally, our mathematical model could be applied to contexts in which the persuasiveness, persuadability, homophily, and partisanship can all be empirically measured. Future work is required in politically-charged contexts and incorporate the potential for “blowback” wherein information across the aisle increases partisanship or ideological conviction for one or both groups (Bail et al., 2018).

## Methods

### Experimental details

**Experiment 1.** 394 Participants were recruited by Cint nearly evenly across the political spectrum. As we are interested in questions about heterogeneity, our sampling paradigm attempted to capture a range of ages, levels of educational attainment and gender representation (Figures 5 and 6).

Participants were initially answered a series of demographic questions. They were then tasked with answering with 20 True/False questions about U.S. State capitals. Participants were either shown the true capital or the largest non-capital city for a random U.S. State. True and False



questions were balanced, and order was randomized for each participant. Participants were presented a with slider and told to indicate their level of confidence in their selected answer. After answering all 20 questions, participants were then provided the same question, with a randomly generated social information statement that read “We asked  $N$  other respondents what they thought was the correct answer and how confident they were. Overall, they believed there was a  $p\%$  chance the answer was true and a  $(1 - p)\%$  chance the answer was false”.  $N$  was chosen uniformly between 1 and 15, and  $p$  was chosen uniformly between 0 and 100. Participants were then given the opportunity to revise their initial estimate.

While Qualtrics was used to present the questions and store responses, questions were generated on the fly using a custom-written Serverless Framework API. Questions about state capitals were used as they are largely non-partisan matters of fact. We used this experimental data to derive an agent-based behavioral model could capture differences across the political spectrum while allowing us to simulate aspects of polarization.

**Experiment 2.** Experiment 2 was identical to experiment 1 from the participants’ perspective. However, this experiment relied on a subset of 10 states identified to be at approximately 50% difficulty (i.e. half the participants were correct). As in experiment 1, participants were shown the correct capital 50% of the time.

Instead of randomly generated social information statements, participants received social information—a guess and a confidence—from randomly chosen set of Experiment 1 participants who had answered the same question. In doing so, the social information that a participant was exposed to matched that seen by the agents in the computational model. Weighting neighbor choices in light of the focal participant’s political leaning and initial correctness allowed us to adjust the degree of homophily and partisanship in these experiments to match those used in our simulation.

**Experiment 3.** For experiment 3, participants ( $N=500$ ) were asked 16 True/False trivia questions focused around comparisons between generally known objects or locations. Examples of these included “Maine is closer to Africa than Africa is to Florida”, and “The Aztec Empire is older than Oxford University.” The compared objects were swapped generating 32 possible questions, half of which were true. These question orders were randomized, to prevent order effects. After completing the trivia questions, participants were given a series of cognitive batteries. Measured attributes included Bullshit Receptivity, Wordsums(verbal intelligence), reliance on Heuristics and Biases, Numeracy, and Need for Cognition and Faith in Intuition (Pennycook et al., 2015; Frederick 2005; Malhotra et al., 2007; Schwartz et al., 1997).

The full survey can be viewed along with the data and code. Experiment 2 and 3 demographics did vary from experiment 1 due to a miscommunication between the experimenters and the third-party sampling service 7. However, we make no direct comparisons between these studies. No participants that completed the experiments were excluded from the analysis. All experiments were conducted in accordance with state and federal regulations and approved by Princeton’s Institutional Review Board.

## Statistical methods

**Overview.** All statistical analysis was conducted using Python and Stan (Carpenter et al., 2017; Stan Development Team 2018). Unless otherwise noted, default parameters in PyStan (v2.19.0.0) were used. Posterior distributions were sampled using Hamiltonian Monte Carlo, and visual inspection of trace plots along with posterior predictive checks were used to evaluate model fit.

## Experiment 1

We estimated the relationship between participant confidence and participant accuracy in Experiment 1 using a hierarchical Bayesian logistic regression (Table 3). A participant’s probability of being correct on a given question was governed by an intercept,  $\alpha_{\text{SUBJ}[i]}$ , an effect of confidence,  $\beta_{\text{SUBJ}[i]}$ , and an estimated question difficulty specific to the state,  $\gamma_{\text{STATE}[i]}$ . The individual subject intercepts and confidence effects were distributed about means values for each political group (Very Liberal to Very Conservative). We placed normal priors on the group mean intercepts, the group mean confidence effects, and the state question difficulties. More formally:

$$\begin{aligned} y_i &\sim \text{Bernoulli}(\Theta_i) \\ \text{logit}(\Theta_i) &= \alpha_{\text{SUBJ}[i]} + \beta_{\text{SUBJ}[i]} C_i + \gamma_{\text{STATE}[i]} \\ \alpha_j &\sim \text{Normal}(\bar{\alpha}_{\text{POLITICS}[j]}, 0.4) \\ \beta_j &\sim \text{Normal}(\bar{\beta}_{\text{POLITICS}[j]}, 0.4) \\ \bar{\alpha}_p &\sim \text{Normal}(0, 0.4), \text{ for } p = 1 \dots 5 \\ \bar{\beta}_p &\sim \text{Normal}(0, 0.4), \text{ for } p = 1 \dots 5 \\ \gamma_s &\sim \text{Normal}(0, 0.4), \text{ for } s = 1 \dots 50 \end{aligned}$$

Priors were selected by prior predictive simulation to ensure that a wide range of plausible relationships were contained in the prior while excluding very extreme relationships between confidence and accuracy.

We likewise estimated overall accuracy for members of political groups using hierarchical Bayesian logistic regression (Table 4). Formally:

$$\begin{aligned}
y_i &\sim \text{Bernoulli}(\Theta_i) \\
\text{logit}(\Theta_i) &= \alpha_{\text{SUBJ}[i]} \\
\alpha_j &\sim \text{Normal}(\bar{\alpha}_{\text{POLITICS}[j]}, 0.4) \\
\bar{\alpha}_p &\sim \text{Normal}(0, 0.4), \text{ for } p = 1 \dots 5
\end{aligned}$$

### Agent-based model

Our agents would be given a position in a network, a political leaning, and an initial answer, either “True” or “False”. To simulate collective wisdom we required a generative model of confidence that differed when agents were correct as well as a model of responsiveness to social information (i.e. whether they can be swayed to change their guess). To accomplish this, we used Bayesian inference to fit parameter estimates to the results from Experiment 1.

**Confidence Generation Model.** For the confidence-generation model, we used a multi-level Zero-One Inflated Beta regression (ZOIB). This was necessary given the slider input, as full and no confidence answers were common (Figure 9). In brief, this mixture model estimates whether a participant will choose to select an extreme example, and if so whether they will select fully confident or unsure. Otherwise, non-extreme answers are modeled as a beta distribution. Each of the  $k = 3$  parameters ( $\alpha_i, \gamma_i, \mu_i$ ) that govern the generation of confidence are modeled with an intercept  $A_{k,j}$  and an effect of being correct  $B_{k,j}$  specific to the  $j^{\text{th}}$  participant. These effects are pooled within the  $p = 5$  political leanings. Formally:

$$\begin{aligned}
y_i &\sim \text{ZOIB}(\alpha_i, \gamma_i, \mu_i, \phi_{\text{POLITICS}[i]}) \\
\text{logit}([\alpha_i, \gamma_i, \mu_i]) &= A_{1:3, \text{SUBJ}[i]} + X[i]B_{1:3, \text{SUBJ}[i]} \\
A_{k,j} &\sim \text{Normal}(\bar{A}_{k, \text{POLITICS}[j]}, 0.5) \\
B_{k,j} &\sim \text{Normal}(\bar{B}_{k, \text{POLITICS}[j]}, 0.5) \\
\bar{A}_{0,p} &\sim \text{Normal}(0, 1) \\
\bar{A}_{1,p} &\sim \text{Normal}(-1, 1) \\
\bar{A}_{2,p} &\sim \text{Normal}(0, 1) \\
\bar{B}_{k,p} &\sim \text{Normal}(0, 1) \\
\phi_p &\sim \text{HalfNormal}(0, 5);
\end{aligned}$$

Priors were chosen using prior predictive simulation to ensure a wide range of plausible distributions of confidence were possible both within and across groups. Notably, the prior for  $\bar{A}_{1,p}$  was set to  $\text{Normal}(-1, 1)$  as it altered the probability that extreme answers (i.e. fully confident, completely unsure) were selected. Had it been set as  $\text{Normal}(0, 1)$ , as with other parameters, the prior would have implied that extreme answers are chosen 50% of the time.

Parameter estimates, based on data from experiment 1, can be found in Table 5. We evaluated the suitability of this

model using posterior predictive checks of both the average confidence and the distributions of confidence across the political spectrum (Figures 8, 9). While the model captured the coarse properties of the participant answers, it failed to capture the tendency of participants to round their confidence to multiples of 5. Nevertheless, as the simulated agents only see aggregated social information we found this limitation acceptable.

**Social influence model.** We next needed a model to determine whether an agent would change their initial guess in the face of conflicting social information. For this, we constructed a multi-level logistic regression in which the probability of switching one’s guess depended on their confidence as well as the extent to which the social information disagreed with their answer. Using the data from experiment 1, we fit this model separately for the conditions in which the participant was initially correct or incorrect. Formally:

$$\begin{aligned}
y_i &\sim \text{Bernoulli}(\Theta_i) \\
\text{logit}(\Theta_i) &= \alpha_{\text{SUBJ}[i]} + \beta_{1, \text{SUBJ}[i]} C_i + \beta_{2, \text{SUBJ}[i]} \text{SI}_i \\
\alpha_j &\sim \text{Normal}(\bar{\alpha}_{\text{POLITICS}[j]}, 1) \\
\beta_{1,j} &\sim \text{Normal}(\bar{\beta}_{1, \text{POLITICS}[j]}, 1) \\
\beta_{2,j} &\sim \text{Normal}(\bar{\beta}_{2, \text{POLITICS}[j]}, 1) \\
\bar{\alpha}_p &\sim \text{Normal}(0, 2), \text{ for } p = 1..5 \\
\bar{\beta}_{1,p} &\sim \text{Normal}(0, 2), \text{ for } p = 1..5 \\
\bar{\beta}_{2,p} &\sim \text{Normal}(0, 2), \text{ for } p = 1..5
\end{aligned}$$

Where  $\beta_1$  and  $\beta_2$  are the effects of confidence (C) and social information (SI) for the  $j^{\text{th}}$  participant, respectively.  $\Theta$  is the probability of changing their guess,  $y$ . Priors were chosen based on prior predictive simulation to ensure a wide range of possibilities within and across groups. These possibilities included the potential for groups to either be highly or not at all susceptible to social influence. As probabilities of switching were not near zero or one, a model cannot accurately predict whether a given participant switches their guess on a specific question. To evaluate model suitability, we used posterior predictive checks to examine the relationship between the predicted and observed rate of switching (Figure 10). Parameter estimates can be found in Tables 6 and 7.

### Simulation study

Using the statistical models described above, we conducted an agent-based simulation to examine the impact of homophily and partisanship on collective wisdom. Partisanship, in this context, was defined by the extent to which liberal agents initially disagreed with conservative agents. This would be impractical to study in a synchronous

experimental context as the questions are not inherently political and it would require a very large pool of participants to construct a partisan social network.

We constructed networks of 500 agents and assigned a portion of them the correct answer. The proportion of agents within each political group was determined at 27% VC, 18% C, 10% M, 18% L, and 27% VL roughly based on existing data from Twitter retweet networks (Brady et al., 2017).

In the nonpartisan condition a proportion of the agents were randomly (corresponding to the difficulty) assigned the correct answer. In the liberal-correct and conservative-correct condition, the same proportion of agents were assigned the correct answer starting with the most extreme political agents on the left and right proportionally.

We then constructed a minimally parameterized network of interaction in which each agent formed between 1 and 7 undirected connections with members on their side of the political spectrum with probability  $h$  and across the aisle with probability  $1 - h$ . As the model only involves one step of social updating based on aggregated social information, the topology of the network and number of connections is unlikely to be impactful when compared to models in which multiple iterations of social interaction occur. Nevertheless, we believed a single round would be justified as changes in collective estimates and wisdom tend to be monotonic and saturating (Becker et al., 2017).

From this arrangement, agents behaved according to parameters drawn from the posterior estimates of the

confidence and social interaction models described above. We then considered a group collectively wise based on a simple majority rule. We evaluated wisdom across 3 levels of homophily (.5, .75, and .98) 20 difficulties (evenly spaced between 40% and 60% initially correct) and considered liberal-correct, conservative-correct and nonpartisan networks. For each condition, we ran 2000 replicates for a total of 360,000 simulations. Simulations were conducted in Python using Numpy, Pandas and the Pandarallel packages. Descriptions of parameters for simulations can be found in the Supplement (Table 1)

## Experiment 2 analysis

The design of experiment 2 involved using the initial answers and political leanings of participants to place them into networks that varied in partisanship or homophily. As a result, selection effects make modeling the accuracy of participant responses in varying conditions intractably confounded. For example, more frequently inaccurate liberal participants would be more likely to end up in a conservative-correct network. This makes it difficult to disentangle, at a participant level, the effect of the selection process and the treatment as certain treatments are over-represented.

To test our predictions while avoiding this issue of selection bias, we aggregated the data from experiment 2 by state (e.g. Virginia, Pennsylvania) and condition (homophily, partisanship). We then modeled the aggregated

**Table 1.** Description of parameters for simulations.

Model	Param.	Desc.	Constraint	Source
Simulations	$d$	Proportion initially correct	$[0, 1]$	Free
—	$h$	Probability a connection formed is with someone from the same group	$[0, 1]$	Free
—	$N$	Size of the group	$[1, \infty]$	Free
Confidence ZOIB	$\alpha$	Probability of making an extreme guess	$[0, 1]$	Fitted
—	$\gamma$	Probability that the extreme guess is 100% confident	$[0, 1]$	Fitted
—	$\mu$	Average confidence when not reported as an extreme.	$(0, 1)$	Fitted
—	$\phi$	Dispersion parameter for $\mu$	$[0, \infty]$	Fitted
Guess Changing	$\alpha$	Intercept, baseline rate of changing one's guess	$[-\infty, \infty]$	Fitted
—	$\beta_1$	Effect of social influence on the probability of changing one's guess	$[-\infty, \infty]$	Fitted
—	$\beta_2$	Effect of social influence on the probability of changing one's guess	$[-\infty, \infty]$	Fitted

**Table 2.** Description of parameters for analytical model.

Param	Description	Constraint
$q_i$	Probability of answering “correctly” in groups 1 or 2 (indicated by subscript)	$[0, 1]$
$a_{ij}$	Probability of correct individuals from group $i$ converting members of group $j$	$[0, 1]$
$b_{ij}$	Probability of incorrect individuals from group $i$ converting members of group $j$	$[0, 1]$
$h$	Probability of interacting with a member of the same group	$[0, 1]$
$\lambda$	Amount of social influence)	$[0, 1]$

accuracy across conditions. To do this, we used a binomial model. More specifically:

$$\begin{aligned} y_i &\sim \text{Binomial}(p_{\text{COND}[i]}, N_{[i]}) \\ \text{logit}(p_{\text{COND}[i]}) &= \alpha_c \\ \alpha_c &\sim \text{Normal}(0, 2), \text{ for } c = 1 \dots 6 \end{aligned}$$

For each of the  $C = 6$  combinations of partisanship and homophily. Prior predictive simulation was used to ensure that a wide range plausible (i.e. near 50%) state-level accuracy values for each condition were possible. Accuracy estimates are shown in Table 8.

### Experiment 3 analysis

For experiment 3, we used a cross-classified hierarchical Bayesian logistic regression to quantify the relationship between confidence and accuracy. This model relied on an overall intercept and effect of accuracy ( $\alpha, \beta_1$ ), as well as a specific intercept and effect for the question and the participant ( $B_2, B_3$ ). Question-specific effects were modeled by the political leaning of the participant. Participant-specific effects were modeled hierarchically considering the political leaning of the participant. Specifically:

$$\begin{aligned} y_i &\sim \text{Binomial}(\Theta_i) \\ \text{logit}(\Theta_i) &= \alpha + \beta_1 \text{Conf}_i + B_{2, \text{QUEST}[i]} \vec{X}_i + B_{3, \text{SUBJ}[i]} \vec{X}_i \\ \alpha &\sim \text{Normal}(0, 1) \\ \beta_1 &\sim \text{Normal}(0, 1) \\ B_2 &= (\text{diag}(\tau_q) \cdot \Omega_q) m_q \\ B_3 &= \gamma u + (\text{diag}(\tau_s) \cdot \Omega_s) m_s \\ \tau_q &= 2.5 * \tan(\tau_{uq}) \\ \tau_s &= 2.5 * \tan(\tau_{us}) \\ \tau_{uq} &\sim \text{Uniform}(0, \pi \div 2) \\ \tau_{us} &\sim \text{Uniform}(0, \pi \div 2) \\ m_{q, 1:Q} &\sim \text{Normal}(0, 1), \text{ for } q = 1 \dots 32 \\ m_{s, 1:S} &\sim \text{Normal}(0, 1), \text{ for } p = 1 \dots 498 \\ \gamma &\sim \text{Normal}(0, 1) \\ \Omega_q &\sim \text{LKJCholesky}(2) \\ \Omega_s &\sim \text{LKJCholesky}(2) \end{aligned}$$

Where  $\vec{X}_i$  is a dummy-coded vector coding the political leaning and confidence of SUBJ<sub>i</sub>,  $u$  are the individual-level predictors (i.e. political leaning, cognitive battery), and  $\gamma$  is a prior on the effect of individual-level predictors. Political leaning and cognitive battery as individual-level predictors were fit separately to avoid confounding one effect by condition on the other. We relied on a non-centered parameterization and modeled the covariance structure of predictors  $B_1$  and  $B_2$  to improve sampling following advice in the Stan manual (Carpenter et al., 2017). Prior predictive simulation was used to ensure that a range of plausible confidence/accuracy relationships were contained in the

prior. Posterior predictive sampling was used to validate model fit.

### Acknowledgements

The authors would like to thank Joanna Sterling, Albert Kao, Andrew Gersick, Benedict Hogan, Mircea Davidescu, Simon Levin, Naomi Leonard, Henry Horn, Damon Centola, Nicholas Christakis and Matt Sosna for feedback on the design and results of the study. They would like to additionally thank Alex Bielen and the developers at Two Bulls for advice on the survey backend, as well as the Princeton Survey Research Center and CINT for assistance in creating and administering the survey.

### Author contributions

D.I.R., I.D.C. and J.B.C. conceived of the study and designed the experiments. C.K.T. and J.B.C. developed the simulations. J.B.C. and D.M. derived the analytical model. All authors were involved in writing and revising the manuscript.


### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Princeton University, The Knight Foundation, National Science Foundation Graduate Research Fellowship (DGE1656466), the UW eScience Institute, CoCCoN, the NSF (IOS-1355061), the ONR (N 00014-09-1-1074 and N00014-14-10635), the ARO (W911NG-11-1-0385 and W911NF14-1-0431), the Struktur-und Innovationsfonds für die Forschung of the State of Baden-Württemberg, and the Max Planck Society.

### Data availability statement

 This article earned Open Data and Open Materials badges through Open Practices Disclosure from the Center for Open Science. Simulation code is available at: <https://github.com/josephbb/Collective-wisdom-in-polarized-groups>. Data and code for all analyses can be found at the Open Science Foundation by clicking here [https://osf.io/jn6wz/?view\\_only=be4fe80f4c51477e8264e463b97f555f](https://osf.io/jn6wz/?view_only=be4fe80f4c51477e8264e463b97f555f).

### ORCID iDs

Joseph B Bak-Coleman  <https://orcid.org/0000-0002-7590-3824>  
Christopher K Tokita  <https://orcid.org/0000-0002-8585-4317>  
Dylan M Morris  <https://orcid.org/0000-0002-3655-406X>  
Iain D Couzin  <https://orcid.org/0000-0001-8556-4558>



## Supplemental material

Supplemental material for this article is available online.

## References

- Almaatouq A, Noriega-Campero A, Alotaibi A, et al. (2020) Adaptive social networks promote the wisdom of crowds. *Proceedings of the National Academy of Sciences*, 117(21): 201917687.
- Bail CA, Argyle LP, Brown TW, et al. (2018) Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences* 115(37): 9216–9221.
- Becker J, Almaatouq A and Horvat A (2020) *Network Structures of Collective Intelligence: The Contingent Benefits of Group Discussion*. Ithaca, New York: ArXiv.
- Becker J, Brackbill D and Centola D (2017) Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences of the United States of America* 114(26): E5070–E5076.
- Becker J, Guilbeault D and Smith N (2021) *The Crowd Classification Problem: Social Dynamics of Binary Choice Accuracy*. Ithaca, NY: Cornell University
- Becker J, Porter E and Centola D (2019) The wisdom of partisan crowds. *Proceedings of the National Academy of Sciences* 166(22): 10717–10722.
- Brady WJ, Wills JA, Jost JT, et al. (2017) Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences* 114(28): 7313–7318.
- Carpenter B, Gelman A, Hoffman M D, et al. (2017) Stan: a probabilistic programming language. *Journal of Statistical Software* 76(1): 1–32.
- Choma BL, Summantry D and Hanoch Y (2019) Right-wing ideology and numeracy: a perception of greater ability, but poorer performance. *Judgement and Decision Making* 14(4): 412–422.
- Condorcet MD (1785) *Essay on the Application of Analysis to the Probability of Majority Decisions*. Paris: Imprimerie Royale.
- Davis-Stober DVB, Budescu DB, Broomel SB, et al. (2014) When is a crowd wise? *Decision* 1(2): 1–4.
- De Polavieja G and Madirolas G (2014) Wisdom of the confident: using social interactions to eliminate the bias in wisdom of the crowds. *Collective Intelligence*. arXiv
- Frederick S (2005) Cognitive reflection and decision making. *Journal of Economic Perspectives* 19(4): 25–42
- Frey V and van de Rijt A (2020) Social influence undermines the wisdom of the crowd in sequential decision making. *Management Science* 67(394): 3713.
- Galam S (2002) Minority opinion spreading in random geometry. *The European Physical Journal B - Condensed Matter and Complex Systems* 2525(44): 403–406.
- Guilbeault D, Becker J and Centola D (2018) Social learning and partisan bias in the interpretation of climate trends. *Proceedings of the National Academy of Sciences of the United States of America* 115(39): 9714–9719.
- Jayles B, Kim H-r, Escobedo R, et al. (2017) How social information can improve estimation accuracy in human groups. *Proceedings of the National Academy of Sciences* 114(47): 201703695.
- Kao AB and Couzin ID (2014) Decision accuracy in complex environments is often maximized by small group sizes decision accuracy in complex environments is often maximized by small group sizes. *Proceedings of the Royal Society B: Biological Sciences* 281(1784): 20133305.
- Kruger J and Dunning D (1999) Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology* 77(6): 1121–1134.
- Laan A, Madirolas G and de Polavieja GG (2017) Rescuing collective wisdom when the average group opinion is wrong. *Frontiers in Robotics and AI* 4: 56.
- Li L, Scaglione A, Swami A, et al. (2013) Consensus, polarization and clustering of opinions in social networks. *IEEE Journal on Selected Areas in Communications* 31(6): 1072–1083.
- Lorenz J and Neumann M (2019) *Opinion Dynamics and Collective Decisions*, Ithaca, NY: Cornell University. 21. DOI: [10.1142/S0219525918020022](https://doi.org/10.1142/S0219525918020022)
- Lorenz J, Rauhut H, Schweitzer F, et al. (2011) How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences* 108(22): 9020–9025.
- Malhotra N, Krosnick J A and Haertel E (2007) *The Psychometric Properties of the GSS Wordsum Vocabulary Test*. National Opinion Research Center: Chicago. Technical report.
- Meisenberg G (2015) Verbal ability as a predictor of political preferences in the United States, 1974–2012. *Intelligence* 50: 135–143.
- Moussaïd M, Kämmer JE, Analytis PP, et al. (2013) Social influence and the collective dynamics of opinion formation. *PLoS ONE* 8(11): e78433.
- Narayanan V, Barash V, Kelly J, et al. (2018) *Polarization, Partisanship and Junk News Consumption over Social Media in the US*. Ithaca, NY: Cornell University
- O'Connor C and Weatherall JO (2018) Scientific polarization. *European Journal for Philosophy of Science* 8(3): 855–875.
- Ortoleva P and Snowberg E (2015a) Are conservatives overconfident? *European Journal of Political Economy* 40: 333–344.
- Ortoleva P and Snowberg E (2015b) Overconfidence in political behavior. *American Economic Review* 105(2): 504–535.
- Pennycook G, Cheyne J A, Barr N, et al. (2015) On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making* 10(6): 549–563.



- Prelec D, Seung H S and McCoy J (2017) A solution to the single-question crowd wisdom problem. *Nature* 541(7638): 532–535.
- Preoiuc-Pietro D, Liu Y, Hopkins DJ et al (2017) Beyond Binary Labels: Political Ideology Prediction of Twitter Users. In: Proceedings of the 55th Conference of the Association for Computational Linguistics, 2017, Vancouver, Canada.
- Schwartz LM, Woloshin S, Black WC, et al. (1997) The role of numeracy in understanding the benefit of screening mammography. *Annals of Internal Medicine* 127(11): 966–972.
- Shi F, Teplitskiy M, Duede E, et al. (2019) The wisdom of polarized crowds. *Nature Human Behaviour* 3: 329–336
- Simmons JP, Nelson LD and Simonsohn U (2011) False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science* 22(11): 1359–1366.
- Stan Development Team (2018) *PyStan: The Python Interface to Stan*.