SOCIAL SCIENCES

Opinion cascades and the unpredictability of partisan polarization

Michael Macy*, Sebastian Deri, Alexander Ruch, Natalie Tong

"Culture wars" involve the puzzling alignment of partisan identity with disparate policy positions, lifestyle choices, and personal morality. Explanations point to ideological divisions, core values, moral emotions, and cognitive hardwiring. Two "multiple worlds" experiments ($n = 4581$) tested an alternative explanation based on the sensitivity of opinion cascades to the initial conditions. Consistent with recent studies, partisan divisions in the influence condition were much larger than in the control group (without influence). The surprise is that bigger divisions indicate less predictability. Emergent positions adopted by Republicans and opposed by Democrats in one experimental "world" had the opposite outcome in other parallel worlds. The unpredictability suggests that what appear to be deep-rooted partisan divisions in our own world may have arisen through a tipping process that might just as easily have tipped the other way. Public awareness of this counter-intuitive possibility has the potential to encourage greater tolerance for opposing opinions.

INTRODUCTION

Escalating political and cultural polarization over the past two decades has attracted growing attention from policy makers and the research community. Concerns include increasingly vitriolic discourse (*1*, *2*) and echo chambers (*3*), with the cumulative potential to erode democratic institutions (*4*). Although the distribution of opinion on most issues remains unimodal (*5*) and polarization is no greater than in earlier periods of extreme political division (*6*), there has been a sharp increase in "party sorting" (*7*, *8*) and partisan antipathy (*9*, *10*).

Polarization is not just the absence of common ground or moderate views. "Partisan polarization" can also entail the division of the population into like-minded groups, with the puzzling alignment of party identity and substantively unrelated public policies, lifestyle preferences, consumer choices, and personal morality (*11*–*13*). "Why on earth," Pinker asks, "should people's beliefs about sex predict their beliefs about the size of the military? What does religion have to do with taxes? Whence the linkage between strict construction of the Constitution and disdain for shocking art?" (*14*).

Survey results suggest a straightforward solution to the puzzle: Partisan polarization across diverse issues reflects the cognitive application of opposing ideological principles (*15*). Other studies report evidence to the contrary, showing that "people endorse whichever position reinforces their connection to others with whom they share important ties" (*16*), a central tenet of the theory of cultural cognition. When those ties are based on party (*17*), political opinions become bumper stickers for "team red" and "team blue," and people sort themselves into opposing teams based on their affinity with the core values celebrated in the bumper stickers (*7*).

Competing explanations focused on ideology versus party identification can nevertheless share the highly plausible assumption that beliefs and opinions diffuse when they resonate—perhaps indirectly and even unconsciously—with deeply rooted psychological predispositions. Disagreements arise as to the relative importance of alternative sources: ideology, identity, core values, "moral emotions" (*18*), and even cognitive hardwiring (*14*).

Once opinions have become hardened and sharply divided, they acquire a mantle of inevitability (*19*). However, models of opinion cascades suggest a radically alternative possibility: The partisan alignment of emergent controversies could be highly arbitrary and unpredictable, due to the sensitivity of path-dependent dynamics to chance events in the initial conditions (*13*, *20*).

Evidence for this explanation emerges from both observation of real-world opinion change and careful examination of opinion cascades in controlled experimental contexts. Zaller's classic model of opinion dynamics shows how political elites (e.g., politicians, pundits, and party leaders) can act as game-changing early movers, especially on unfamiliar issues about which a highly susceptible general public has yet to acquire sufficient information to form independent opinions (*21*). For example, Zaller points out how Nixon's uncommon implementation of wage and price controls in 1971 triggered a marked about-face among Republican rank and file. The case study presages equally unexpected contemporary reversals of long-standing partisan positions on issues like free trade, NATO leadership, and legalization of marijuana.

The unpredictability of cascade outcomes was most notably demonstrated in a unique online experiment involving artificial cultural markets in multiple worlds (*22*). The "Music Lab" results were unexpected: Previously unfamiliar "Indie" songs that became highly popular in one "world" were, in some cases, among the least popular in another. If megahits like *Star Wars* or *Harry Potter* might have easily been flops (*23*), could opinion cascades also explain the emergence of partisan divisions that are nonetheless idiosyncratic?

To find out, we used the "multiple worlds" experimental paradigm to test the possibility that opinion cascades can also generate unpredictable political alignments in which the advocates on an issue might have instead been the opposition but for the luck of the draw in the positions taken by early movers. In two experiments, equal numbers of Democrat and Republican participants located in the United States were recruited over a four-week period using Prime Panels, an online market research platform (www.turkprime.com/Service/PrimePanels). In each experiment, participants were first asked with which party they identify and how strongly. They were then randomly assigned to 10 parallel worlds of uniform size and administered a survey with up to 20 randomly ordered political

Cornell University, Ithaca, NY 14850, USA.
*Corresponding author. Email: mwm14@cornell.edu

and cultural statements (see table S2). In 8 of 10 worlds (the influence condition), participants could see which party was more likely to agree with an item, while in the other two worlds (the independence condition), they could not. In the influence condition, the participant's own agreement was then used, in turn, to update the relative support of each party displayed to the next participant in that same world. Participants only knew about their own world and did not even know that there were other worlds. (Additional details on the experimental designs are included in the Supplementary Materials.)

The experiments were designed to test hypotheses about the sensitivity of cascades to early movers who weigh in before most other people on a newly emerging controversy. We therefore pretested the survey items to minimize preexisting opinion or partisan predisposition (the political equivalent of Music Lab's use of unfamiliar Indie music).

Participants' opinions were solicited by asking "Now we would like to know your own individual opinion. As a [Democrat/Republican] do you agree or disagree with this statement?"

The responses were used to test two hypotheses about cascade dynamics in the alignment of opinions:

H1: Partisan differences will be larger in the worlds with social influence compared to the independence condition.

H2: The alignment of issues will be less predictable between worlds in the influence condition than in the independence condition.

Both hypotheses were formulated before data collection and were included in a 2017 proposal to the National Science Foundation that funded the experiment.

METHODS

The hypotheses were tested in two related experiments, one framed as a survey and the other as a game. (Details of the differences between the two designs are included in the Supplementary Materials). The survey version asked participants to respond to 20 items on unfamiliar issues "that might become future controversies." The game version challenged participants to correctly predict the position their party would take on each issue. Both versions precluded interactions among the participants.

Extensive pilot testing demonstrated five difficulties creating ecologically valid social influence under controlled laboratory conditions. First, a blank slate is a rare occurrence for potentially controversial political issues. None of the items in the pretest were entirely free of pre-existing opinion or party alignment. Second, social influence in survey research is discouraged by normative expectations for independent judgment. Music Lab recruited from a music download website where users were accustomed to choosing popular songs to listen to for the first time. Following the crowd is normative on music download sites but not in opinion surveys. Third, neither experiment involved participant interactions. Social influence would likely have been much stronger had we conducted the experiment as a "chat room" where participants might experience social pressure for in-group conformity and out-group hostility. However, enabling participant interaction would greatly constrain the number of issues and participants. Fourth, similar to Music Lab, the experiment excluded exposure to news reports, social movement mobilizations, or advertising campaigns that reinforce cascades with repeated top-down application of persuasive messages, symbolic meanings, and collective identities. Last, downloading music is highly engaging, but participants in online surveys may click through mechanically because of boredom or fatigue.

To strengthen social influence, we used blue and red fonts for the survey item and the agree and disagree buttons to reinforce party identification and to provide an easy reminder of the information we provided about the party alignment on the item among previous participants. Depending on the party with the higher proportion in agreement on an item, participants in the influence condition saw the statement displayed in either blue (Democrats more likely to agree), red (Republicans more likely), or purple (for equal proportions). Participants in the independence condition (and first movers in the influence condition) always saw the issue in a purple font. In soliciting the participant's opinion, we worded the question to activate party identification by prefacing with "As a Democrat/Republican," but also asking for "your individual opinion."

RESULTS

The results were highly consistent between the two experiments, and we begin with the results of the simpler survey version ($n = 2271$).

Experiment 1: "The Future Controversies" survey

Figure 1A reports the effects of social influence on agreement with previous participants. The figure reports the proportion of responses that conformed with the majority of previous participants from the same party, broken down by party and strength of party identification (four groups in each treatment condition). The proportions are visualized relative to chance.

In the absence of information about previous participants, agreement could nevertheless reflect shared beliefs among those from the same party that differ from the beliefs of the other party. This level of agreement with previous participants provides a baseline against which to measure susceptibility to social influence when information about previous participants is provided. In the independence condition, only Republicans (regardless of the strength of party identity) were above chance in aligning with their party (0.55, $P < 0.001$), indicating stronger ideological predispositions than among participants from the other party. In the influence condition, all four groups were above chance. Across the 20 items, the proportion of participants' responses that were aligned with their party in the influence condition was 0.62, compared to 0.51 in the independence condition ($d = 0.11$, $P < 0.001$ adjusting for multiple tests and non-independence of repeated measures). These results confirm previous research showing a strong effect of party identification on political opinion (*17*).

Social influence, in turn, led to stronger partisan alignments in the influence condition (H1). Partisan alignment in Fig. 1B is measured from −1 to 1 as the mean proportion of Democrats who agree on an item in a specific world minus the proportion of Republicans who agree, with the absolute values averaged over 20 items and all worlds in the independence or influence condition. In the independence condition, differences in the items' intrinsic appeal led to partisan alignments above chance (0.1 and 0.08 among those with and without strong party identification; $P < 0.001$). However, the effects of social influence were larger still. Among participants with and without strong party identification, partisan alignment was 0.29 and 0.21 ($P < 0.001$) in the influence condition. Social influence more than doubled the level of partisan alignment ($d = 0.19$ 0.13, $P < 0.001$) and additional analyses show no attenuation of the treatment effects over time.

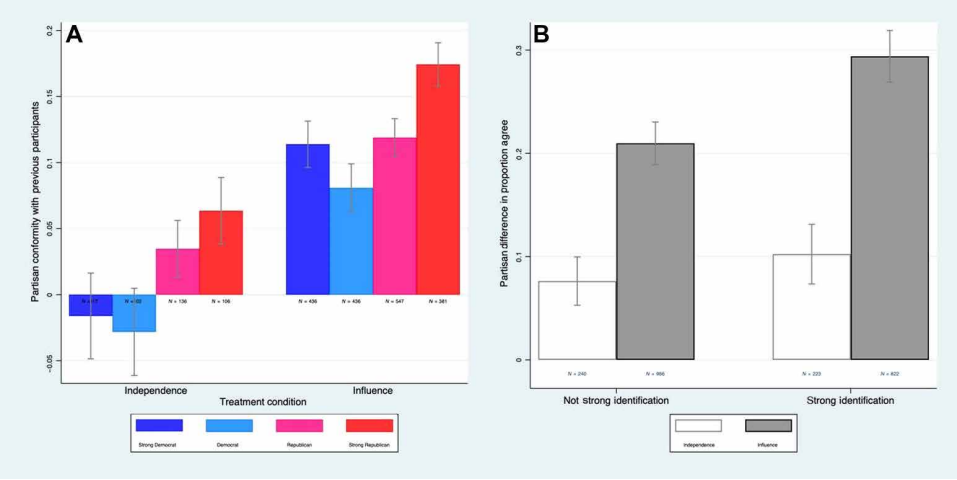

Fig. 1. The effects of social influence on partisan conformity and partisan alignment. (**A**) The effect of social influence on partisan conformity with previous participants. The figure reports the proportion of responses that conformed with the majority of previous participants from the same party, broken down by party (blue for Democrats and red for Republicans), strength of party identification (color intensity), and treatment condition. The proportions are scaled relative to an even chance of conforming ($Y = 0$), and the four categories are exclusive. First movers and other participants who saw equal support from both parties were not included. Confidence intervals of 95% were calculated using a bootstrapping procedure with 100 replications for each random draw with $n-1$. Conformity was almost at chance among Democrats in the independence condition but not among Republicans, indicating greater intrinsic Republican ideological appeal/aversion. In the influence condition, Republicans remained more likely to conform than were Democrats, especially those with a strong party identity. (**B**) The effect of social influence on partisan alignment. Influence from previous participants, in turn, led to a stronger partisan alignment in the influence condition compared to the independence condition, especially among participants with strong party identification. Alignment is measured as the mean proportion of Democrats who agree on an item in a specific world minus the proportion of Republicans who agree, with the absolute values averaged over 20 items and all worlds in the independence or influence condition. Regardless of the strength of party identification, the differences of opinion between the parties were nearly three times as large as in the influence condition (0.21 and 0.29) compared to the independence condition (0.08 and 0.11).

Figure 2 provides a more detailed view of partisan alignments in each of the 10 worlds. The *Y* axis measures partisan alignment as the proportion of Democrats who supported the issue (in blue) minus the proportion of Republicans (in red), preserving the sign (unlike Fig. 1B that reports the absolute value). The horizontal dashed lines demarcate the boundary of statistical significance (at $P < 0.001$, with Bonferroni adjustment for 20 tests in each world). In worlds 1 and 2 (the independence condition), no issues became politically aligned above chance. In the influence condition, substantively unrelated issues became linked into party-specific profiles, much like the "culture wars" in our own world. For example, in world 3, partisan majorities took opposite sides on six issues: Republican majorities agreed and Democrats opposed the items on great books (15), obedience (19), and gas engines (20), while Democrats agreed and Republicans opposed items on robot lawyers (3), licensed jurors (7), and group loyalty (13). Recall Pinker's puzzlement about "the linkage between strict construction of the Constitution and disdain for shocking art" (*14*). Figure 2 provides an alternative to ideology, moral emotions, or cognitive hardwiring as the explanation: Social influence that aligns diverse opinions with political party (*17*).

If partisan alignments reflected differences in intrinsic ideological affinity, we should expect only chance variation from one world to another. That is not what we found. For example, Fig. 2 shows how a Democratic majority endorsed item 17 (on student character) in world 7, while Republicans were opposed. However, in four other worlds, the issue was supported by a Republican majority, while Democrats opposed. Democrats favored (and Republicans opposed) an issue about genetic information in two worlds, but the sides were reversed in two others. On 13 of 20 issues, the party more likely to agree in one world was more likely to disagree in another.

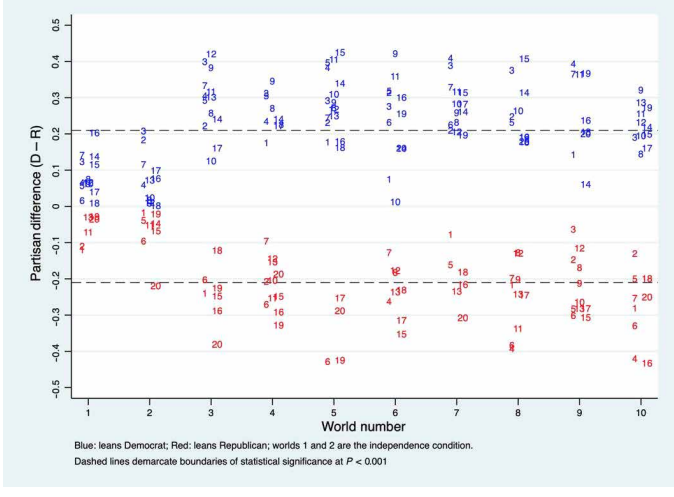

Fig. 2. Partisan alignment by world and by survey item. The *Y* axis measures partisan alignment as the proportion of Democrats who supported an issue (in blue) minus the proportion of Republicans (in red). (Issue content by issue number is available in table S2.) Dashed lines demarcate the boundaries of statistical significance at $P < 0.001$, with adjustment for 20 tests in each world. The parties did not differ significantly on any issues in the independence worlds (1 and 2). In the influence worlds (*3*–*10*), substantively unrelated issues became linked into party-specific profiles, much like the culture wars in our own world. For example, in world 3, Republican majorities agreed and Democrats opposed the items on great books (15), obedience (19), and gas engines (20), while Democrats agreed and Republicans opposed items on robot lawyers (3), licensed jurors (7), and group loyalty (13).

Figure 3 shows how the partisan alignment of issues was more unpredictable in the influence condition (H2). We measured predictability in the independence condition using the same

bootstrapping procedure that was used in Music Lab: randomly sorting all 463 participants in the independence condition into two newly constituted worlds with equal expected size and repeating the procedure 1000 times (see section S1, B.3 for details). We then compared the partisan alignments in each of the 1000 pairs of randomly drawn worlds. For the influence condition, we compared the 28 pairs of eight influence worlds. For each pair, in each treatment, we tested four pairwise indicators of unpredictability: the mean pairwise difference in partisan alignment and the probability of a change in the sign of partisan alignment (i.e., the other party becomes more likely to agree), as well as a change in the majority position in one or both parties (see table S3 for illustrations of the measures).

The results confirm the higher probability of a pairwise change in the influence condition for each of the four measures. Averaged over the 560 comparisons of partisan alignment on 20 items in 28 world pairs, the mean pairwise difference was 0.28 in the influence condition, ranging from 0.1 to 0.42 and with an SD of 0.07, compared to only 0.006 in the independence condition. These between-world differences in partisan alignment, in turn, led to qualitative differences between worlds in the party more likely to favor a given item, as reported in Fig. 3. Changes in the party more likely to agree were at chance (49%), and while reversals of majority support were less likely (40% for one party and 20% for two), these events remained far more unpredictable than in the independence condition.

As an additional test of predictability, we measured how well partisan alignments are predicted by the partisan alignment in the independence condition (Fig. 4A) and by the partisan alignment of the early movers (Fig. 4B). (Panel A also labels the item numbers in blue above the *X* axis, as well as the partisan alignment below the *X* axis.) The quadratic estimate indicates that intrinsic ideological appeal (as observed in the independence worlds) can bias cascades in a preferred direction. In panel B, the *X* axis is the alignment among the first 10 movers on a given item in each of the eight influence worlds, and the *Y* axis reports the alignment of all who followed. The results show that the alignment of early movers on an item is a better predictor of later alignments ($r = 0.67$) than knowing the outcome for that item in the independence condition ($r = 0.18$). (The *Y* axis in panel B excludes the first 10 movers; we also repeated the analysis excluding the first 100 movers, with nearly identical results; $r = 0.65$.)

Issues also varied widely across worlds in their alignment with other issues. In the most extreme case, participants in the influence condition who agreed with item 4 (on foreign tax policy) were more likely to agree on seven other items in some worlds (the seven pairwise correlations were positive and significant at $P < 0.001$). However, in other worlds, supporters of item 4 were less likely to agree on each of these same seven items.

More generally, knowing that two items were correlated in one world was a poor predictor that this pair would be similarly correlated in another world. Of 190 pairs of 20 items, 129 were statistically significant ($P < 0.001$) in one or more influence worlds, again showing how social influence causes substantively unrelated issues to align. However, the alignments were highly idiosyncratic. On average, the 190 pairwise correlations in any one of the eight influence worlds explained only 3% of the variance in the correlations in one of the other seven influence worlds. For example, the significant negative correlation in world 3 ($r = -0.23$) between item 1 (banning cryptocurrencies) and item 11 (life will be better in 100 years) nicely predicts the correlation between those same two items in world 10 ($r = -0.18$) but not the positive correlation in world 8 ($r = 0.20$).

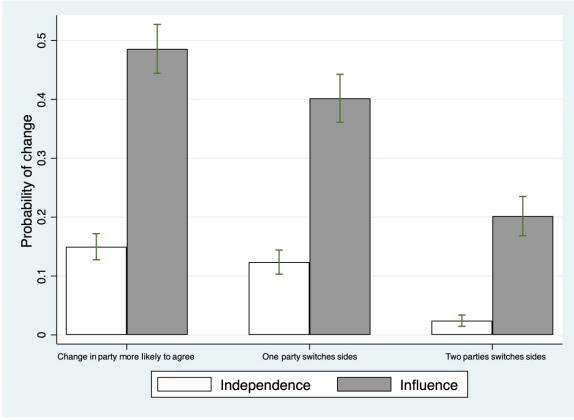

Fig. 3. Effect of social influence on the unpredictability of partisan alignment. The *Y* axis measures unpredictability using three indicators: a change in the party more likely to agree and whether one or both parties switched sides on an issue. The measures for each item were averaged over 28 pairs of eight worlds in the influence condition ($N = 560$) and over 1000 random splits of the 463 participants in the independence condition into two worlds of equal size. The results challenge the intuition that large partisan disagreements reflect deeper underlying ideological differences in the intrinsic appeal of political opinions. Recall Fig. 1B, showing that partisan alignments were nearly three times stronger in the influence condition compared to the independence condition. Figure 3 shows that the alignments are nevertheless more unpredictable as well, due to the sensitivity of cascades to chance variation in the opinions of early movers.

In sum, the results suggest that intrinsic appeal needs to be sufficiently strong relative to social influence, or its effects can be tipped the other way by chance variation in a small number of early movers. These reversals between worlds defy the intuition that large differences between the parties are likely to be highly robust (*24*). The intuition is correct if the differences reflect intrinsic appeal but not when they are the outcome of a cascade.

Experiment 2: The Prediction Game

A parallel "Prediction Game" experiment ($N = 2310$) demonstrates the robustness of these results. The items, treatment conditions, and worlds were identical, but the framing changed from an opinion survey (the "Future Controversies" study) to a Prediction Game in which participants competed to win $100 by correctly predicting the party more likely to agree with the opinion in each of the study items. The game format was intended to avoid the normative obligation in a survey to provide an independent judgment.

Following the display of the item, participants in the influence condition were told the partisan views of previous participants, using identical wording as in the Future Controversies version. However, before giving their own opinion, participants were presented with two intervening tasks. First, participants were asked to pick the most plausible of three possible explanations for their party's alignment on each issue, ideology, history, or popularity. The explanation task was intended to encourage participants to attend to and internalize the influence signal. Next, participants were asked to predict the party that was more likely to agree. The prediction task was intended to focus the participant's attention on the position the party was likely to take instead of the position the party should take. Following the explanation and prediction tasks, participants were then asked to give their individual opinions, as in the Future Controversies version of the experiment.

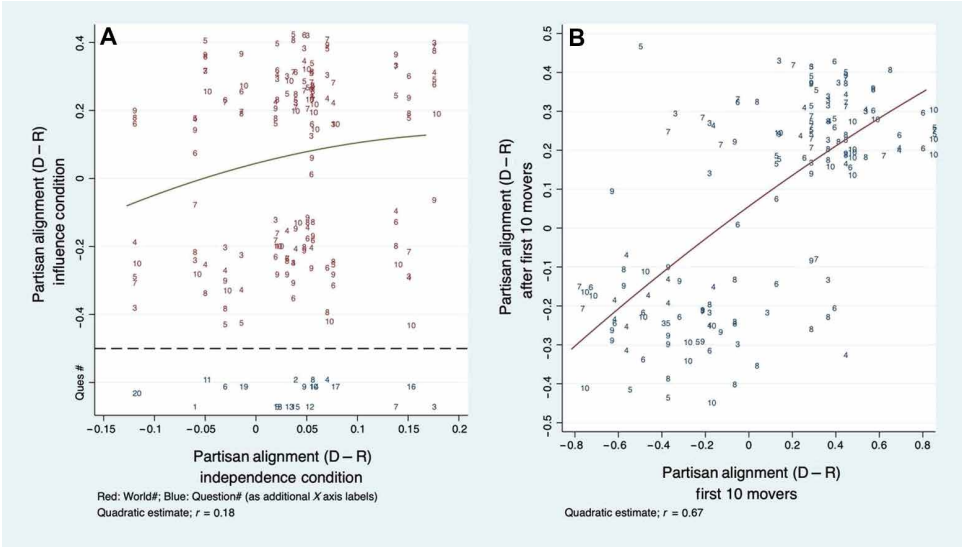

Fig. 4. Effects of intrinsic appeal and early movers on partisan alignment. (**A**) The partisan alignment on each item in each of the eight influence worlds (*Y* axis) is predicted by the alignment in the independence condition (*X* axis). The *X* axis is also labeled by corresponding item numbers (in blue, vertically jittered above the *X* axis). Alignment is measured as the proportion of Democrats who agree on an item minus the proportion of Republicans. The slope of the quadratic estimate indicates that intrinsic ideological attraction to an issue can bias the cascade in a predictable direction ($r = 0.18$). (**B**) The partisan alignment on each item in each of the eight independence worlds among the first 10 movers (*X* axis) predicts the alignment among those who followed the first 10 movers (*Y* axis). Comparison of the two panels shows how the influence of the first 10 movers ($r = 0.67$) can outweigh the effects of intrinsic ideological appeal ($r = 0.18$).

Both experiments used 10 multiple worlds, with 2 in the independence condition and 8 in the influence condition. Both versions also exposed participants in the influence condition to information about previous participants, using identical wording, and both asked participants to respond to the same 20 survey items. However, the Future Controversies version was structured as a survey and only asked participants to respond to each item by indicating whether they agreed. In contrast, the second experiment was structured as a Prediction Game in which participants competed for a $100 prize (a monetary incentive not used in experiment 1) by trying to predict which party would be more likely to agree with the item, based on an independent survey (that is, participants knew that the winning prediction would be judged using different data, not the data from the game in which they were participating).

Following a practice question (not used in the data analysis), participants were presented with the same survey items used in the Future Controversies experiment, in a random order. Because of the additional time and effort to perform the explanation and prediction tasks, participants in the Prediction Game were randomly assigned only 15 of the 20 items (instead of all 20). All 20 items were used, but with fewer responses on each item than if everyone were assigned all 20 items. In the influence condition, each of the 15 survey items was presented with the same screen used in the Future Controversies study, using the same red and blue fonts to designate the political parties and the same summary of the responses of previous participants.

The prediction task motivated participants with a monetary incentive (a chance to win $100) to make a correct prediction, for which they could consider the responses of previous participants, although they did not know the number and could therefore not use statistical generalization. They could also consider their previous assessment of the reasons the parties might agree with the item from the explanation task, which focuses attention on the intrinsic partisan appeal instead of the opinions of previous participants.

Although we had expected the monetary incentive and prediction task to strengthen social influence, it appears to have had the opposite effect, perhaps due to the separation of the agreement decision from the information about the behavior of previous participants by two intervening tasks (the explanation and prediction tasks). Analysis of the chain of decision (explanation to prediction to opinion) identified the weak link in the chain: the prediction task. Weaker influence was due to "contrarian" predictions that ran counter to those of previous participants. The contrarians then expressed agreement in line with their predictions, not the partisan preferences of previous participants. A total of 38% of all predictions in the influence condition were contrarian, and of the 1853 participants in the influence condition, 1634 (88.2%) made at least one contrarian prediction. Contrarian behavior did not vary by party or by whether the participant agreed with the item. The most plausible explanation appears to be attention shift. Contrarian predictions may have been influenced by the preceding explanation task, which focused the participant's attention on the intrinsic appeal of the item to one party instead of the trending opinion of previous participants. The explanation and prediction tasks appear to have shifted participants' attention away from information about the partisan alignment of previous participants and toward the sources of intrinsic appeal encapsulated in the three explanatory options.

Nevertheless, social influence was still strong enough to produce cascades with partisan divisions and unpredictable outcomes. The results in figs. S1 to S4 are qualitatively similar to those from the Future Controversies version (Figs. 1 to 4). The largest difference is fig. S4, where the correlation between partisan alignment in the independence and influence conditions was much stronger ($r = .69$) than in the Future Controversies experiment ($r = 0.43$). As a result, partisan alignments in the independence condition were a stronger

predictor of alignments in the influence condition than were the first 10 movers ($r = 0.42$, compare to 0.67 in Future Controversies).

The robustness of the results across the two experiments shows that the results do not depend on a particular design, while the differences show that unpredictability depends on the strength of social influence. The Prediction Game attenuated social influence and strengthened consideration of the item's intrinsic appeal, and this, in turn, led to less unpredictability than observed in the survey-based design. Comparisons of the results of the two versions of the experiment recommend a balanced conclusion: Participants' choices reflect the relative strength of intrinsic appeal and social influence. Political polarization becomes unpredictable when social influence is stronger than intrinsic appeal.

van de Rijt's recent reanalysis of the Music Lab data showed that the influence worlds were converging toward the worlds without influence (*25*). We therefore tested for convergence over time in partisan alignments between the first, middle, and last third of participants. The results show no indication of convergence between worlds or between treatment groups in either version of the experiment. The absence of change over time means that neither the parties nor the treatment groups were converging. Nevertheless, we cannot rule out the possibility that convergence might occur were the experiment to continue indefinitely.

Given the concerns raised by van de Rijt, we identified the mechanism responsible for convergence in Music Lab: the attenuation of the stimulus. Each time a participant downloaded a currently unpopular song, they nudged that song up the "charts" for the next participant, thereby compounding the error by increasing the probability that the next participant would also download that same song. The solution is to insulate the stimulus from the accumulation of error. Instead of providing participants with quantitative information about the distribution of party support on each item (e.g., a bar chart), we provided only qualitative information as to which party was more favorable. So long as participants were more likely than not to follow the norm, the cascade could only be reversed by a chance string of consecutive deviants, and the longer the cascade runs, the longer the string needed to reverse it. This design thus satisfies a key requirement for testing the hypothesized sensitivity of cascades to early movers. Once the issue is tipped to one party or the other, it becomes increasingly difficult to tip it back. This design also relaxes the need to test for long-term convergence by recruiting thousands of participants to each world. Last, the qualitative information has persuasive ecological validity. In natural settings, it is much easier for people to tell whether an issue is backed by Democrats or Republicans than to know precisely how popular the issue might be with each party. Moreover, if a Democrat happens to adopt a Republican position on an issue, then she may increase the chances that her network neighbors will deviate as well, but she does not affect the probabilities of all members of the population, as would happen if we were to give every participant in her world a bar chart that changed with every participant's response.

DISCUSSION

These two experiments support methodological and substantive conclusions. The differences between worlds provide an important methodological reminder: Correlates of opinion (such as party identification) that do not take cascade dynamics into account should invite the same skepticism as zero-order correlations without statistical controls for causal priors. Opinion cascades in the underlying population may be accurately reflected in the correlated responses to the survey, which, in turn, are likely to be a good predictor of another sample from that same population—but not necessarily of another sample from a second population, even one that is otherwise identical to the first.

Our study also addresses scientific debates about the causes of political polarization. To be clear, we do not try to explain the vicissitudes or time scale of opinion dynamics that has been observed in real historical settings, nor do we question the idea that strong religious, cultural, and ideological commitments can inform opinions on hot-button issues like reproductive rights. They obviously do. Our study tests an alternative explanation for newly emerging political and cultural divisions. Previous scholarship on polarization has debated the relative importance of ideology, core values, and party identification. Our results suggest that partisan alignments across substantively diverse issues do not necessarily reflect intrinsic preferences but may indicate instead the outcome of cascade dynamics that might have tipped in a different direction due to chance variation in the positions taken by early movers. Public awareness of this counterintuitive possibility has the potential to encourage greater tolerance for alternative opinions.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/5/8/eaax0754/DC1

Section S1. Materials and Methods: Descriptive details

Section S2. The Prediction Game

Table S1. The distribution of participants over worlds, by party and strength of party identification.

Table S2. Survey items, with susceptibility to social influence.

Table S3. Illustrations of the three pairwise measures of unpredictability used in Figs. 2 and 3.

Fig. S1. The effects of social influence on partisan conformity and partisan alignment.

Fig. S2. Partisan alignment by world and by survey item.

Fig. S3. Effect of social influence on predictability of partisan alignment.

Fig. S4. Effects of intrinsic appeal and early movers on partisan alignment.

REFERENCES AND NOTES

1. S. Iyengar, S. J. Westwood, Fear and loathing across party lines: New evidence on group polarization. *Am. J. Pol. Sci.* **59**, 690–707 (2015).
2. S. Iyengar, G. Sood, Y. Lelkes, Affect, not ideology: A social identity perspective on polarization. *Publ. Opin. Quart.* **76**, 405–431 (2012).
3. S. Flaxman, S. Goel, J. M. Rao, Filter bubbles, echo chambers, and online news consumption. *Publ. Opin. Quart.* **80**, 298–320 (2016).
4. C. Sunstein, *Republic: Divided Democracy in the Age of Social Media* (Princeton Univ. Press, 2009).
5. M. P. Fiorina, S. J. Abrams, Political polarization in the American public. *Annu. Rev. Pol. Sci.* **11**, 563–588 (2008).
6. D. Brady, H. Han, in *Red and Blue Nation?: Characteristics and Causes of America's Polarized Politics*, P. Nivola, D. Brady, Eds. (Brookings Press, 2006), chap. 3, pp. 119–151.
7. L. Mason, "I disrespectfully agree": The differential effects of partisan sorting on social and issue polarization. *Am. J. Pol. Sci.* **59**, 128–145 (2015).
8. M. Fiorina, *Unstable Majorities: Polarization, Party Sorting, and Political Stalemate* (Hoover Institution Press, 2017).
9. J. Gramlich, Far more Americans say there are strong conflicts between partisans than between other groups in society. Pew Research Center, 19 December 2017.
10. A. I. Abramowitz, S. W. Webster, Negative partisanship: Why Americans dislike parties but behave like rabid partisans. *Pol. Psychol.* **39**, 119–135 (2018).
11. M. Dimock, Political Polarization in the American Public. Pew Research Center, 12 June 2014.
12. J. Kiley, In polarized era, fewer Americans hold a mix of conservative and liberal views. Pew Research Center, 23 October 2017.
13. D. DellaPosta, Y. Shi, M. Macy, Why do liberals drink lattes? *Am. J. Soc.* **120**, 1473–1511 (2015).
14. S. Pinker, *The Blank Slate: The Modern Denial of Human Nature* (Viking, 2003).

15. S. W. Webster, A. I. Abramowitz, The ideological foundations of affective polarization in the U.S. electorate. *Am. Pol. Res.* **45**, 621–647 (2017).
16. D. Kahan, Fixing the communications failure. *Nature* **463**, 296–297 (2010).
17. G. L. Cohen, Party over policy: The dominating impact of group influence on political beliefs. *J. Pers. Soc. Psy.* **85**, 808–822 (2003).
18. J. Haidt, *The Righteous Mind* (Vintage, 2013).
19. D. Watts, *Everything is Obvious (Once You Know the Answer)* (Crown, 2011).
20. D. Baldassarri, P. Bearman, Dynamics of political polarization. *Am. Soc. Rev.* **72**, 784–781 (2007).
21. J. Zaller, *The Nature and Origins of Mass Opinion* (Cambridge, 1992).
22. M. J. Salganik, P. S. Dodds, D. J. Watts, Experimental study of inequality and unpredictability in an artificial cultural market. *Science* **311**, 854–856 (2006).
23. C. Sunstein, *The World According to Star Wars* (Harper Collins, 2016).
24. S. Goel, A. Anderson, J. Hofman, D. J. Watts, The structural virality of online diffusion. *Manage. Sci.* **62**, 180–196 (2015).
25. A. van de Rijt, Self-correcting dynamics in social influence processes. *Am. J. Soc.* **124**, 1468–1495 (2019).

Acknowledgments: We thank A. Goldberg, A. V. Loon, R. Willer, and members of the Social Dynamics Lab for their contributions and especially J. Qu who coded the experimental design. IRB guidelines were followed for research with human subjects. **Funding:** The study was funded by the U.S. National Science Foundation (SES 1756822 and GRFP 2017241254), with additional support from the Korean National Research Foundation (NRF-2016S1A3A2925033), the DARPA "Ground Truth Challenge" (146626), and the Minerva Initiative (FA9550-15-1-0162).
Author contributions: All authors participated equally in designing the experiment, pre- and pilot testing the design, replicating and double-checking the results, and editing and revising the manuscript. M.M. conceived the study, ran the online experiment, and conducted most of the data analysis and writing. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 20 February 2019
Accepted 19 July 2019
Published 28 August 2019
10.1126/sciadv.aax0754

Citation: M. Macy, S. Deri, A. Ruch, N. Tong, Opinion cascades and the unpredictability of partisan polarization. *Sci. Adv.* **5**, eaax0754 (2019).

ScienceAdvances

Opinion cascades and the unpredictability of partisan polarization

Michael Macy, Sebastian Deri, Alexander Ruch and Natalie Tong

Sci Adv **5** (8), eaax0754.
DOI: 10.1126/sciadv.aax0754