SPECIAL FEATURE: PERSPECTIVE

Polarization, diversity, and democratic robustness

Jenna Bednar[a,b,c,1]

Edited by Simon Levin, Ecology and Evolutionary Biology, Princeton University, Princeton, NJ; received August 19, 2021; accepted November 2, 2021

In the Madisonian Constitution, fragmented and overlapping institutions of authority are supposed to manage democracy's innate rivalry, channeling competition to serve the public interest. This system of safeguards makes democracy more robust: capable of withstanding and, if need be, adapting to challenges posed by a changing problem environment. In this essay, I suggest why affective polarization poses a special threat to democratic robustness. While most scholars hypothesize that polarization's dangers are that it leads to bimodality and extremism, I highlight a third hypothesized effect: Polarization reduces interest and information diversity in the political system. To be effective, democracy's safeguards rely upon interest diversity, but Madison took that diversity for granted. Unique among democracy's safeguards, federalism builds in a repository for diversity; its structure enables differences between national- and state-expressed interests, even within the same party. This diversity can be democracy hindering, as the United States' history with racially discriminatory politics painfully makes clear, but it can also serve as a reservoir of interest and information dispersion that could protect democracy by restoring the possibility that cross-cutting cleavages emerge.

robust systems | democratic decline | federalism | polarization

Surging polarization challenges the United States' democracy. In the past, polarization's main threat seemed to be gridlock as parties refused to compromise (1). Recently, the nature of polarization has changed; political scientists label it affective polarization (2–4), where the out-group is not just a rival but an enemy. Affective polarization is sufficient to create policy polarization but a darker form, with motivations blind to policy rationale. When the other group is detested, policy position taking is based on out-group opposition rather than a careful review of a policy's efficacy. The United States' democratic status has been downgraded, and political scientists are pointing to polarization as the cause (5, 6). Democracy is meant to channel the competition of rivals for the common good. Why would it be vulnerable to the rivalry created by polarization?

In this essay, I open by describing the constitutional safeguards in the United States' democracy and the theory of how they contribute to democratic robustness. I then examine three processes that affective polarization is hypothesized to set in motion and their subsequent effect on democracy: entrenched rivalry by distanced parties, runaway process that spirals beyond elite control, and compression of interest and issue diversity. Of the three hypothesized effects, the last is most dangerous to democracy because it destroys the information and interest diversity that is essential for democracy's safeguards to work. Without the protection of its safeguards, democracy is vulnerable to autocracy. In institutionalizing diversity, federalism plays a special—although not fail-safe—role in democracy's resilience. It also has its own capacity to pervert democracy.

Democracy's Defense System

The Madisonian Constitution. As we have come to mythologize James Madison and the "Madisonian Constitution," the genius of the Constitution is that it does not leave human nature to sort itself out but instead, creates institutions to structure human interaction. In a democracy, the people rule, but they can make choices against their common interests, especially long-term interests. The Madisonian Constitution is designed to accomplish two tasks, diametrically opposed: to create an "energetic" government

[a]Ford School of Public Policy, University of Michigan, Ann Arbor, MI 48109; [b]Department of Political Science, University of Michigan, Ann Arbor, MI 48109; and [c]Santa Fe Institute, Santa Fe, NM 87501

Author contributions: J.B. designed research, performed research, and wrote the paper.

The author declares no competing interest.

This article is a PNAS Direct Submission.

[1]Email: jbednar@umich.edu.

Published December 6, 2021.

strong enough to fend off foreign invaders and domestic insurrection, while at the same time, prevent the rise of a tyrant who serves narrow interests rather than the broader public good.

Key to making democracy work is the design of government. Under the Articles of Confederation, Congress was feeble; it could only make requests of the states for money or action. While the Constitution endowed the national government with true authorities, its true strength is not the expanded power of government but its restraints. Government power is checked by a system of safeguards: structural fragmentation, electoral accountability, judicial oversight, and federalism. Each of these safeguards is imperfect and incomplete—the word "safeguard" is misleading because it implies that any safeguard might be sufficient or guaranteed to work. Just as a physical guardrail might not stop the heaviest trucks from tumbling over a road's edge, democracy's safeguards might fail. As I will describe below, they can complement one another to improve their performance.

Of these four types, we associate the structural safeguards most tightly with Madison: bicameral legislature, separation of powers, and an independent judiciary. The separation severs chains of command, making tyranny harder to assemble. The overlapping power creates a system of mutual accountability. Federalism would provide a "double security," acting as a further fragmentation of authority and point of resistance. Behind the fragmented apparatus, the people held the reins. The hope was that this system-within-system design would be sufficient to arrest the rise of a despot while creating a government that could serve the collective needs of a burgeoning, well-endowed new country.

System Robustness. The founders sought to establish a system of governance that would last in perpetuity. With time comes variability—in problem scope, in resource availability, and in stakeholder desires. Given the need to adapt to changing conditions, successful governments are not stable as much as robust; they are able to continue to function, perhaps with adaptation, despite perturbations (7, 8).

In a democratic system, the "perturbations" that threaten the integrity of popular self-rule include the temporary shock of corrupt leadership, permanent realignments in the global economy or interstate alliances, and changing public conceptions of citizenship and equality. Governmental authority should be able to address current needs and change to meet future needs. The safeguards need to be able to uphold the existing authority boundaries while managing their adaptation. How well institutions handle the unanticipated determines the democratic system's robustness.

Ideally, robust democratic system design relies on distributed, simple safeguards rather than a single, complex, catchall structure. Safeguards should be appropriate to their environment; they may be judges who review constitutionality of laws, presidents who object to perceived congressional overreach and veto legislation, or voters who react to the performance of policy. In this way, each institution has distinct and limited capacities, but as a composite, they jointly safeguard democracy.

The architecture of robust design includes three features: diversity, modularity, and redundancy (8–10). The safeguard architecture can be too permissive, tolerating behavior that should be stopped. To insure against failure, robust systems have built-in redundancy, backups that maintain functional continuity despite component failure. Additionally, for failures that cannot be averted, modularization contains failures while the rest of the system learns and adjusts.

Diversity—of interests, of issues—is a source of new ideas in the system. If a system of safeguards is too rigid, there is no exploration of ideas and practices that might provide a vision forward when a system needs to adapt. Additionally, if it lacks interest dispersion, there is no potential for cross-cutting cleavages to develop. The success of the architecture is measured by its capacity to manage the perturbations, avoiding maladaptation or system collapse.

The Robustness of the Federal Union. The founders did not have the benefit of a theory of robustness when they designed the Constitution, but the design—using complexity to their advantage, incorporating diversity, redundancy, and modularity—often makes it seem as if they had a copy of Levin's *Fragile Dominion* (9) in hand while drafting. Different structural safeguards—separation of powers, bicameralism, independent judiciary, staggered elections, and term lengths—all employ redundancy; these institutions of governance overlap and duplicate one another, with the mutual accountability that comes from that overlapping oversight. The differing selection mechanisms—different electorates, sequential election timing, indirect appointments (state or presidential), and yes, even the Electoral College—were devised to draw in a diversity of perspectives and guard against the passions that could lead an electorate to misunderstand its best actions. Should one component be overtaken by passion—say, the House of Representatives—then the Senate, with the longer view afforded by longer term lengths, or the president, with a national perspective, could resist the House's mischief.

While they did not articulate an appreciation for the benefits of modularity (that came later, beautifully, by Justice Brandeis) (11), in *Federalist* 46 Madison made clear that federalism was a strength of the new American system; should the institutions of the national government somehow pursue action counter to the general interest, despite its redundancy, the states could issue a "signal of general alarm" that would alert the voting public to the national government's overreach (12).

So, the system of safeguards was designed to channel rivalry, keeping political opponents on one playing field. It also was intentionally structured to embrace diversity of inputs; the boundaries of acceptable behavior would be set and patrolled by a diverse, redundant, and modular set of institutions (8).

The Problem of Polarization

We intuitively think of affective polarization as entrenched rivalry. To understand how affective polarization poses a unique threat to democracy, we can isolate three of its hypothesized effects: social divisiveness, runaway processes that elude elite control, and suppression of issue and interest diversity.

Effect 1: Homophily and Aversion Lead to Bimodal Distributions. To explain how and why affective polarization leads to bimodal partisan distributions, two forces are frequently invoked: homophily and out-group aversion. Homophily, the preference to be near others who are like you, explains much locational clustering (13). A second force captures the push that comes from avoiding those who are dissimilar. This push force—called out-group derogation, aversion, or in political science, affective polarization—is less benign than homophily; it represents a judgment, a distaste

for those who differ. Coupled with emotion, these judgments become passionate, leading to hatred. It would seem to be the more dangerous of the two forces (2, 14, 15).

Additionally, a common class of models includes a dynamic of conformity, where agents become more like one another as they sort (16). Like homophily, agents are drawn to others who share traits, but with conformity, agents change features of themselves to become even more like others. If in addition, people seek internal consistency and their internal moral compass or set of beliefs and principles is synonymous with those expressed by a political party, then these two forces are mutually reinforcing, accelerating the process of full partisan conformity and depersonalization (17).

Most group dynamics models include both our desire to assimilate and our desire to differentiate, and we refer to them as attraction–repulsion models (ARMs). Affective polarization creates two distanced partisan rivals, more sharply divided than other forms of polarization.

Effect 2: Elites Lose Control of Polarization. Political scientists hold competing views about whether democracy is lost from the top—when a despot dupes an unwitting public by dismantling democratic institutions—or from the bottom—when the public comes to value autocratic rule over democracy.* In any case, elites certainly stir up partisanship to serve their private interests.

To explain elite-driven polarization, ref. 20 assumes that voters satisfice, supporting the candidate closest to their ideal. As parties strategically reposition in response to one another, they reduce their inclusiveness—their ideological scope—while they shift farther apart from one another. The process is slow and easily disrupted. Elites might speed up the process by manipulating mass beliefs: for example, when elites convince their followers that it is treacherous—even treasonous—to deviate from the party line in any way.

Psychologists have found that one of the most effective ways to stimulate a blindly loyal following is not through homophily but through an aversion based on emotion and fear. Emotion speeds and amplifies group dynamics; fear, especially anxiety about death, converts mild in-group bias to out-group hatred. When people are terrified, they cluster more tightly within their in-group. Social norms against injuring others fade away as people become willing to harm members of the out-group, whom they perceive as posing an existential threat (14).

Elites may be able to stir up polarization, but they can lose control of it because the process feeds on itself, eluding control. In the ARM (21), the feedback between attraction and repulsion forces quickly becomes a runaway process. Ref. 22 uses the same push–pull dynamic combined with exogenous shocks to demonstrate that polarization processes have unpredictable dynamics including irreversibility; as the process passes a tipping point, polarization becomes entrenched. The ARM-based model in ref. 23 begins with a change in public opinion; an asymmetry in the feedback processes of self-reinforcement and reflexive partisanship leads to their conclusion that the Republican Party has already reached this threshold of irreversibility and cannot be restored from within. Additionally, the rise of social media influencers—a rival elite type—not only exacerbates polarization (24) but obstructs the moderate elite's ability to connect with the voters.

When elites catalyze fear, they goad a benign human tendency to assimilate and differentiate into a runaway process, building to the point where polarization is irreversible. Counterintuitively, affective polarization may also lead to extremism; empirical studies suggest that deliberation among the like-minded will push the group's position to the extreme of the members' initial positions (25, 26). Neither individual commitment to democratic norms nor elite moderation can restore bipartisanship. In fact, evidence points to the contrary; in a polarized world, voters prioritize their partisan interests over democratic principles, and electoral safeguards fail to protect democracy (27).

Effect 3: Polarization Undermines Interest and Issue Diversity. Polarization paralyzes governments, but entrenched rivalry is not sufficient to destroy democracy; the safeguards were designed to manage this competition. A third consequence of affective polarization is particularly threatening: the compression of the information space, including the suppression of diverse interests and issues.

Polarization does more than divide a population; it simplifies it. When people conform to one another, they let go of their differences. Recall the ARMs I describe above. When they also include trait conformity, not only are groups becoming more sharply divided, but they do so by becoming more like those in their in-group. At the system level, information is lost. When agents conform, the dimensionality of the issue or identity space is reduced.

Compression of the information space can happen for a number of reasons. In the model of Kawakatsu et al. (28) of polarization leading to issue compression, individuals maximize their utility through pairwise cooperation. As polarization increases, it becomes easier to coordinate with other individuals (28). As social identification with a party grows, political preferences are shaped more by peer copartisans than elites (29). These conformist pressures are internally enforced. Psychologists have long identified the "black sheep effect," where peers punish in-group deviants more strongly than out-group deviants (30). Additionally, moral outrage spreads quickly online, feeding social conflict and out-group dehumanization (31). In conjunction with an emotionally driven process, the loss of elite control, and the expansion of social media's influence, with affective polarization, in-group policing becomes more powerful, leading to a tighter clustering of the in-group.

Polarization with out-group aversion can lead to the underadoption of beneficial innovations. If one group adopts an innovation of universal benefit, the other group will avoid it (15). Call it the cut off your nose to spite your face result—or an explanation for why mask wearing and now, vaccinations have differential uptake between Democrats and Republicans.

Copying others can be adaptive if learning is costly (32). However, maintaining information in the system can be helpful in the long run if the problem environment changes. Too much conformity can leave a population without the fresh ideas it needs to confront new challenges, and the highest-performing organizations will embrace some nonconformity, particularly when they operate in volatile environments (33). Additionally, diversity opens the possibility of cross-cutting cleavages and bridges that are so helpful for compromise.

*An example of an elite-driven argument is in refs. 6 and 18, and pushback emphasizing the mass-driven process is in ref. 19. Ref. 1 cautions that the underlying mechanism may be impossible to untangle.

Democracy is like a social information processor; at best, it takes the breadth of human interests and ideas and turns them into useful policy to promote social welfare. Democracy relies on the free flow of diverse information to be innovative and productive (34, 35).

Diversity is an asset, and polarization with conformity burns that asset. Additionally, as we will see, it also undermines the supports girding our democracy against a slide toward autocracy.

Polarization and Robustness

Diversity's Vulnerability. Each of the three effects I hypothesize above would cause democracy to underperform. However, underperformance is not the same as leaving democracy unshielded from the rise of an autocrat, the consequence that safeguards are meant to prevent. Of the three, for reasons I develop next, information compression most undermines democracy's institutional safeguards.

The safeguards designed to make democracy robust rely on diversity—of interests, of ideas, of preferences, and of perspectives. Democracy is not like a basketball game where third-party neutral referees ensure that players follow the rules; in a democracy, all safeguards are internal to the system and have biases. Each safeguard, whether structural, judicial, political, or electoral, is an amalgamation of interests. For safeguards to protect democracy from tyranny, the inputs into that amalgamation must remain diverse.

Robust democracies structure their safeguards to diversify input. For example, the structural safeguards—the separation of powers—fragment authority and draw representation at staggered intervals from different aggregations: district, state, and national (through a state-weighted composite). The court system, made up of judges with their own biases and political concerns, takes pains to build a culture of objectivity but with its own double security of varied partisan input in judicial appointments to encourage balancing. Voters also often fail to hold democratic principles above their private or partisan interests (6), and so, as an aggregate—an electorate—they are most effective if the breadth of their interests finds a voice in the ballot box. So it goes with the full slate of safeguards. It is the diversity that make them effective veto players.

Safeguards are more than independent veto players; they also complement one another (8). Safeguards are staffed by humans with their own biases and missteps; they are imperfect. If one fails to stop governmental abuse, then another is available, as when an unpopular court decision leads to a social uprising. As a robust system, for them to be effectively redundant—that is, fulfilling functionality when another fails—they must fail for different reasons. As safeguards, they will fail if all are captured by a single interest and if they all align with the same tightly circumscribed party. The effectiveness of the system of safeguards to bolster democratic robustness depends crucially on diversity.

Madison took that diversity for granted. He believed that it was innate to human nature as well as a product of commercial activity (36). While he worried about the coagulation of public opinion and the likelihood that people would support falsehoods, Madison believed that the institutions were sufficient. He never considered that a nation as sprawling as ours would ever lack for diversity.

In a democratic system, diversity substitutes for neutrality. In compressing information and interests, polarization eliminates the diverse sources of failure. Without diversity, the safeguards become aligned, failing as one to prevent the rise of an autocrat.

Remedy: Multipolarity? If polarization suppresses diversity, how might the system restore it? Rather than fight against polarization, perhaps one might encourage multipolarity. The underlying dynamics of attraction and repulsion can support multiple poles; Axelrod's model (16) of cultural diffusion (homophily with conformity) produces stable equilibrium distributions containing more than two groups. Could multipolarity reduce polarization's destructiveness?

While we do not have an explicit model of multipolar ARM with conformity, we can run a thought experiment drawing upon the dynamics described by the bipolar ARMs. Suppose one subpopulation dislikes another subpopulation; that is, suppose we have out-group aversion. With just two poles, aversion pushes those who dislike one pole toward the only other pole. With multipolarity, a push away from one pole is not equivalent to a pull toward another pole. Those with out-group aversion may not necessarily end up at the same pole but could be split between two or more poles, themselves a minority, their repulsion simmering beneath the surface while other interests and concerns shape the identity of their pole. Unless a new pole emerges that taps this common out-group aversion, they will not be united.

Empirical evidence of supporters of Donald Trump appears to back this conjecture. In 2011, data show that out-group aversion—a measurable "animus toward minority groups"—existed within both Democratic and Republican Parties (37). Panel data show that Donald Trump's emergence catalyzed their union; out-group aversion, rather than homophily, was the initial force.

It might have stayed at that, with a new third pole around Trump, except that the American political system is set up to favor two national parties (38). While third parties have from time to time made a small showing, the animus that Trump tapped was more significant than the percentages that previous third-party candidates like Ralph Nader or Ross Perot could win. After the new pole around Trump formed, elites could spur loyalty-based conformity to drive out rivals and keep the group coherent. The Republican moderates found the wave unstoppable, consistent with the runaway models in this volume. Issues that had been Republican concerns prior to 2016 fell to the side: free trade, strong alliances, small government, immigrant amnesty, and respect for the Constitution and governing institutions. The loss of diversity from Trumpism is the loss of the bubbling up of ideas within the Republican Party as well as the partisan-bridging potential of a more complex issue space.

Almost by definition, multipolarity supports a more complex information space and greater diversity. It is possible that multipolarity would reduce some of the extreme outcomes; even better, from the viewpoint of democracy, is that multipolarity would encourage the cross-cutting cleavages and other forms of complex alliance formation and information exchange that support compromise. Many democracies have electoral systems that support multiple parties, and that multipolarity may be one reason that there appears to be less affective polarization in other democratic systems.[†] However, in the United States,

[†]Comparative data on affective polarization suggest that the United States has higher out-party dislike but not higher in-party liking than other countries. The data stop at 2017 (39).

multiparty-created multipolarity appears out of reach absent reform to the United States' electoral system, and that reform is unlikely to be supported by the partisans who benefit from the current system.

Remedy: Federalism? There is another possible route to diversification: federalism. Of democracy's safeguards, only federalism has a source of diversity built into it. Each state has a distinct economy, geography, resource base, and demographic profile. Each has a history and trajectory, a set of challenges and self-defined goals. No state duplicates another. These unique interests and needs mean that, by definition, federalism brings diversity to the democratic system.

When Madison referred to federalism as democracy's double security, he envisioned, and later exercised, the capacity of states to stand up as rivals to challenge national government overreach, extending the political game board. This coordinated resistance across the states is what he hoped might develop with the Virginia and Kentucky Resolutions.

Federalism maintains diversity in the system, but its capacity is not well represented by maps of red and blue states or by red states vs. a blue national government. We are at a high point in within-state partisan variation (40), and while national parties have grown deft at coercing state parties to adopt their policy preferences (41), policy implemented within a state will develop state-specific contours.

These contours are evidence of federalism's capacity to expose cracks within a partisan monolith. Even at a moment when the Republican Party tightly controls its membership, in the past year we have seen state-level Republicans bucking the party's position: secretaries of state and election board members who followed procedure rather than bending the rules to change election outcomes; Republican governors who coordinated pandemic response with Democratic governors in neighboring states; Republican governors who expressed regret for banning mask mandates; and at least one Republican governor who reached out to invite resettlement of Afghani immigrants in his state, Utah. While these public acts are not at present a measurable trend, they might signal a loosening of partisan conformity.

Federalism's ability to fracture partisan blocks can lead to terrible outcomes; it enabled antidemocratic racist policies to persist in the Southern US states long after white-only policies were ended elsewhere (42). Federalism has too many qualitative variations to be easily plugged in as a dummy variable to make a straightforward prediction like "federations are more robust than unitary systems." That said, if the federal structure is robust—that is, if the subnational governments are able to exercise some meaningful independence from the national one—then it will always provide one thing: a source—a possibility—of diversity, for better or for worse.

Federalism may be the seed that begins a turn toward democratic backsliding by enabling authoritarianism to take hold in a state. However, in providing an extra layer of diversity, redundancy, and modularization, it may also be a repository of democracy or where democracy begins anew. Federalism guarantees a living democracy, not a static one. Federalism's inescapable churn provides opportunities for new ideas, information, and interests to constantly bubble up. It restores dimensionality to the political space, creating the potential for cross-cutting cleavages to develop. It is far from a guarantee of a flourishing society, but in the face of partisan conformity that contributes to democratic decay, it does institutionalize hope.

Data Availability. There are no data underlying this work.

Acknowledgments

I acknowledge the Andrew Mellon Foundation for sponsoring my sabbatical at the Center for Advanced Study in the Behavioral Sciences and the University of Michigan for its support of my research leave.

1 N. McCarty, *Polarization: What Everyone Needs to Know* (Oxford University Press, 2019).
2 S. Iyengar, Y. Lelkes, M. Levendusky, N. Malhotra, S. J. Westwood, The origins and consequences of affective polarization in the United States. *Annu. Rev. Polit. Sci.* **22**, 129–146 (2019).
3 C. Perrings, M. Hechter, R. Mamada, National polarization and international agreements. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102145118 (2021).
4 A. J. Stewart, J. B. Plotkin, N. McCarty, Inequality, identity, and partisanship: How redistribution can stem the tide of mass polarization. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102140118 (2021).
5 P. Pierson, E. Schickler, Madison's constitution under stress: A developmental analysis of political polarization. *Annu. Rev. Polit. Sci.* **23**, 37–58 (2020).
6 M. W. Svolik, Polarization versus democracy. *J. Democracy* **30**, 20–32 (2019).
7 E. Jen, Stable or robust? What's the difference? *Complexity* **8**, 12–18 (2003).
8 J. Bednar, *The Robust Federation: Principles of Design* (Cambridge University Press, 2009).
9 S. Levin, *Fragile Dominion: Complexity and the Commons* (Basic Books, 1999).
10 S. A. Levin, "The architecture of robustness" in *Global Challenges, Governance, and Complexity: Applications and Frontiers*, V. Galaz, Ed. (Edward Elgar, 2019), 16–23.
11 New State Ice Co. v. Liebmann, 285 U.S. 262 (1932).
12 J. Madison, The Federalist Papers: No. 46. The Avalon Project, Yale Law School. https://avalon.law.yale.edu/18th_century/fed46.asp. Accessed 23 July 2021.
13 T. C. Schelling, *Micromotives and Macrobehavior* (Norton, 1978).
14 M. Hewstone, M. Rubin, H. Willis, Intergroup bias. *Annu. Rev. Psychol.* **53**, 575–604 (2002).
15 P. E. Smaldino, M. A. Janssen, V. Hillis, J. Bednar, Adoption as a social marker: Innovation diffusion with outgroup aversion. *J. Math. Sociol.* **41**, 26–45 (2017).
16 R. Axelrod, The dissemination of culture: A model with local convergence and global polarization. *J. Conflict Resolut.* **41**, 203–226 (1997).
17 J. Bednar, A. Bramson, A. Jones-Rooy, S. Page, Emergent cultural signatures and persistent diversity: A model of conformity and consistency. *Rationality Soc.* **22**, 407–444 (2010).
18 S. Haggard, R. Kaufman, *Backsliding: Democratic Regress in the Contemporary World* (Cambridge University Press, 2021).
19 P. Norris, Voters against democracy. *Foreign Aff.* **100**, 174–178 (2021).
20 V. C. Yang, D. M. Abrams, G. Kernell, A. E. Motter, Why are U.S. parties so polarized? A "satisficing" dynamical model. *SIAM Rev.* **62**, 646–657 (2020).
21 R. Axelrod, J. J. Daymude, S. Forrest, Preventing extreme polarization of political attitudes. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102139118 (2021).
22 M. W. Macy, M. Ma, D. R. Tabin, J. Gao, B. K. Szymanski, Polarization and tipping points. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102144118 (2021).
23 N. E. Leonard, K. Lipsitz, A. Bizyaeva, A. Franci, Y. Lelkes, The nonlinear feedback dynamics of asymmetric political polarization. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102149118 (2021).
24 J. Flamino *et al.*, Emergence of new influencer elites challenges traditional polarization analysis. arXiv [Preprint] (2021). https://arxiv.org/pdf/2111.02505.pdf (Accessed 3 November 2021).

25 C. R. Sunstein, Deliberative trouble? Why groups go to extremes. *Yale Law J.* **110**, 71–119 (2000).
26 D. Schkade *et al.*, What happened on deliberation day? *Calif. Law Rev.* **95**, 915 (2007).
27 M. Graham, M. W. Svolik, Democracy in America? Partisanship, polarization, and the robustness of support for democracy in the United States. *Am. Polit. Sci. Rev.* **114**, 392–409 (2020).
28 M. Kawakatsu, Y. Lelkes, S. A. Levin, C. E. Tarnita, Interindividual cooperation mediated by partisanship complicates Madison's cure for "mischiefs of faction." *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102148118 (2021).
29 B. Toff, E. Suhay, Partisan conformity, social identity, and the formation of policy preferences. *Int. J. Public Opin. Res.* **31**, 349–367 (2019).
30 J. M. Marques, D. Paez, The 'black sheep effect': Social categorization, rejection of ingroup deviates and perception of group variability. *Eur. Rev. Soc. Psychol.* **5**, 37–68 (2011).
31 M. J. Crockett, Moral outrage in the digital age. *Nat. Hum. Behav.* **1**, 769–771 (2017).
32 R. Boyd, P. J. Richerson, *Culture and the Evolutionary Process* (University of Chicago Press, 1988).
33 J. G. March, Exploration and exploitation in organizational learning. *Organ. Sci.* **2**, 71–87 (1991).
34 S. E. Page, *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies* (Princeton University Press, 2007).
35 H. Farrell, B. Schneier, Democracy's Dilemma. *Boston Review*, 15 May 2019. https://bostonreview.net/forum-henry-farrell-bruce-schneier-democracys-dilemma. Accessed 6 April 2021.
36 J. N. Rakove, *A Politician Thinking: The Creative Mind of James Madison* (University of Oklahoma Press, 2017).
37 L. Mason, J. Wronski, J. V. Kane, Activating animus: The uniquely social roots of Trump support. *Am. Polit. Sci. Rev.* **115**, 1508–1516 (2021).
38 L. Drutman, *Breaking the Two-Party Doom Loop: The Case for Multiparty Democracy in America* (Oxford University Press, 2020).
39 N. Gidron, J. Adams, W. Horne, *American Affective Polarization in Comparative Perspective* (Elements in American Politics, Cambridge University Press, Cambridge, United Kingdom, 2020).
40 D. Darmofal, R. Strickler, *Demography, Politics, and Partisan Polarization in the United States, 1828–2016* (Springer International Publishing, 2019).
41 J. M. Grumbach, From backwaters to major policymakers: Policy polarization in the states, 1970–2014. *Perspect. Polit.* **16**, 416–435 (2018).
42 R. Mickey, *Paths out of Dixie: The Democratization of Authoritarian Enclaves in America's Deep South, 1944-1972* (Princeton University Press, 2015).