Political Polarization and Democratic Erosion: Mapping the Institutional Design Gap in Contemporary Scholarship

Banat Singh

Independent Researcher

banat.singh.research@gmail.com

ORCID - https://orcid.org/0009-0005-2942-0839

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

Literature Review

The literature on political polarization has undergone a fundamental reconceptualization over the past two decades, yet it remains plagued by a persistent and consequential blind spot that limits both theoretical understanding and practical solutions. Scholars have migrated from early frameworks treating polarization as policy divergence among elites, through increasingly sophisticated psychological accounts of affective division, toward an emerging third-generation understanding that begins—only begins—to grapple with political incentives and structural constraints. Yet even this latest iteration stops short of the conceptual apparatus necessary to explain why identical levels of emotional antipathy produce radically different democratic outcomes across ostensibly similar societies, why individual-level interventions consistently fail to produce collective-level change, and why the causal mechanisms linking institutional design to mass polarization remain stubbornly underspecified in the literature.

The problem is not that scholars have failed to notice institutions. It is rather that institutional explanations have remained largely additive—institutions are appended to psychological and informational models as an additional variable in a regression equation—rather than constitutive of the phenomenon itself. This review argues that the field has developed around a fundamental category error: the treatment of polarization as a phenomenon amenable to individual-level causal explanation, when the actual mechanisms producing it operate at the systemic level of institutional design, knowledge architecture, and feedback loop dynamics. This gap in understanding represents more than a theoretical lacuna; it has direct and consequential effects for how we conceptualize solutions and design interventions. If polarization is fundamentally an individual-level problem of attitude formation and emotional response, then individual-level interventions—correcting misperceptions, facilitating intergroup contact, priming common identity—represent plausible and potentially effective remedies. But if polarization is fundamentally a system-level problem produced by coupled institutional mechanisms operating simultaneously, then such interventions cannot

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

work not because they are poorly designed or implemented, but because they target the wrong level of analysis entirely. This review maps that conceptual gap systematically and argues for a reframing grounded in institutional design methodology as an analytical approach.

Addressing this gap requires integrating three distinct but previously siloed research traditions: comparative institutional analysis that specifies mechanisms in detail rather than merely correlating institutions with outcomes; systems dynamics theory that models feedback loops and emergent properties of complex systems; and design methodology from architecture and human-centered design that treats institutions as artifacts produced through deliberate choices and amenable to deliberate redesign. The resulting framework would not solve polarization—wicked problems are not solved but only managed and negotiated—but it would provide the conceptual apparatus to ask genuinely better questions and propose more plausible approaches to democratic governance in persistently polarized contexts. By integrating these approaches, scholars can move beyond individual-level explanations toward systemic understanding that accounts for how institutional choices structure political behavior across entire populations.

The Limits of First and Second Generation Frameworks

The intellectual history of polarization scholarship reveals a pattern of successive paradigm migrations, each addressing anomalies that the previous framework could not accommodate. The first generation, crystallized in seminal work by Poole and Rosenthal and developed further by scholars like McCarty, identified congressional polarization as ideological distance between legislative voting records.[1] This framework proved empirically robust for its narrow domain: members of Congress demonstrably have moved further apart ideologically since the 1970s, with little overlap between parties on recorded votes.[2] The explanatory mechanisms offered were sophisticated and multifaceted—declining competitive districts, changing campaign finance incentives, declining informal social interaction among legislators across party lines—and the policy implications seemed clear and actionable: reform electoral systems to increase competition,

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

restore earmarks and budget mechanisms as tools for coalition-building, and create space for bipartisan socialization within legislative institutions.

The strength of this first-generation approach was its methodological rigor and empirical data-driven focus on measurable outcomes. The DW-NOMINATE scores that measured congressional ideology became the gold standard for polarization measurement across the field, and the steady upward trend in polarization over decades seemed to establish a clear and undeniable empirical fact. The framework was parsimonious: polarization was ideological distance, it could be measured precisely through voting patterns, and its causes could be subjected to quantitative analysis. Yet this parsimony came at a cost. The approach treated polarization as a property of voting behavior alone, which is only a narrow slice of actual political behavior. It said nothing about what citizens actually believed, how political parties mobilized voters to support their positions, or what consequences polarization had for democratic governance beyond legislative gridlock and delay.

Yet this framework encountered a stubborn and persistent obstacle that forced scholars to reconceptualize their fundamental understanding of the phenomenon itself. The polarization observable in Congress failed to scale cleanly and predictably to the mass public. Multiple studies found significant policy overlap among ordinary voters on ostensibly polarizing issues like guns and abortion. Citizens held ideologically inconsistent views, with Americans regularly adopting positions from both left and right sides of the political spectrum on different issues. Moderates existed in surprising abundance—perhaps as much as a quarter or more of the electorate held mixed ideological positions depending on the issue domain. The standard explanation—that elite polarization was sorting the mass public, with polarized voters then selecting polarized representatives—proved only partially correct. Careful analysis of candidate pools revealed that more ideologically extreme candidates had been running for office since the 1980s, suggesting supply-side sorting among political entrepreneurs rather than demand-side sorting among voters.[4] This finding was theoretically significant because it suggested that elite polarization might be

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

somewhat endogenous to political institutions rather than simply reflecting underlying voter preferences. More troublingly still, even when scholars demonstrated that ordinary Americans held cross-cutting policy preferences across domains, those same citizens reported intense emotional dislike for members of the opposing party. The puzzle was not that people disagreed on policy—or at least, not primarily that—but that their emotions seemed entirely orthogonal to their actual policy positions. How could citizens share policy preferences yet harbor profound and persistent hostility? How could they simultaneously agree on specific policies but view the opposing party as fundamentally threatening to the nation?

This anomaly catalyzed the second generation framework and intellectual shift. By the early 2010s, scholars working at the intersection of political psychology and comparative politics proposed that polarization should be fundamentally reconceptualized not as policy distance but as affective polarization—the emotional dislike, social distance, and identity-based hostility felt toward opposing partisans.[5] This move proved clarifying in multiple ways simultaneously. Affective polarization had risen steadily since the 1980s, with some measures suggesting accelerating growth from the 2000s onward. It appeared in mass publics even where policy disagreement was objectively modest. It could be measured directly through feeling thermometers (asking respondents how warmly they felt toward the other party), questions about demographic stereotypes of opposing partisans, and willingness to have children marry across party lines. Importantly, affective polarization varied substantially across democracies in ways that pure information or inequality measures could not explain. Countries with similar media landscapes, similar levels of inequality, similar educational distributions, showed dramatically different levels of affective polarization depending on their institutional structures. The mechanism seemed psychological rather than informational: people were not divided because they lacked information about their actual agreement, but because they held distorted beliefs about the other side, saw their own group as fundamentally threatened, and experienced the opposing camp not as merely wrong but as morally deficient or even dangerous.

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

The second-generation framework generated an enormous and increasingly sophisticated literature on underlying mechanisms. Scholars identified identity-based reasoning, where partisanship functioned as a master identity that colored all other identities—religious identity, racial identity, geographic identity, educational identity.[7] They documented in-group affinity and out-group threat perception, showing how social categorization processes produced hostility independent of any actual conflict of interest or resource competition. They traced the role of elite rhetoric carefully, asking whether dehumanizing language and affective appeals by politicians and media figures were creating an environment hostile to cross-partisan engagement. And crucially, they developed and tested interventions systematically: correcting misperceptions about the other side, priming common identity, bringing people together across difference under structured conditions with skilled facilitation.

These interventions worked—at least in laboratory settings and short-term experimental contexts with self-selected participants. When researchers exposed participants to information showing that the other party was less wealthy, less evangelical, less urban than people believed, affective polarization decreased measurably. When they primed American identity or showed videos of cross-partisan agreement on common values, emotional dislike declined significantly. When they brought carefully selected groups together for moderated dialogue with trained facilitators, participants reported warmth toward individual out-partisans and reduced prejudice afterward. The meta-analysis of intergroup contact theory suggested that, under proper conditions, bridging programs could reduce the emotional distance that seemed to be tearing democracies apart.[10] The logic was compelling and theoretically parsimonious: if affective polarization was the primary driver of democratic dysfunction, and if affective polarization could be reduced through psychological and relational interventions, then scaling up such interventions should improve democratic health substantially.

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

Yet something crucial was missing from this framework, though it took several years and extensive empirical testing for the field to recognize it clearly. As Rachel Kleinfeld's comprehensive assessment of the polarization literature reveals with unusual clarity, interventions that successfully reduced affective polarization in laboratory settings produced zero detectable change in antidemocratic attitudes, support for democratic norm-breaking, or actual political violence.[11] Participants whose emotional dislike of the other party diminished by 20 percent still voted identically in subsequent elections. They did not become more willing to accept election results they disliked, or less willing to support their party if candidates employed antidemocratic means to win. The emotional warmth toward individual out-partisans did not translate into behavioral change in the real political world. Most troublingly, when researchers examined the actual temporal and demographic patterns of political violence in the United States systematically, the data proved inconsistent with an affective polarization explanation. Affective polarization had been rising steadily since the 1980s; political violence spiked sharply and suddenly after 2016. Affective polarization was nearly symmetrical across parties; political violence was overwhelmingly concentrated on the right. Affective polarization was rising fastest among Americans over sixty-five, who consume primarily television and talk radio; yet if affective polarization were the driver of political violence, we would expect the fastest growth among this group to correlate with the most violence, yet the data showed no such pattern. The empirical foundation of the second-generation model—that affective polarization is the primary cause of democratic erosion and violence—began to crumble when examined rigorously.

The Institutional Incentives Problem: Where Current Literature Stalls

The third generation of scholarship, still emergent and intellectually incomplete, has begun to grapple with the limitations of psychological and informational frameworks by reintroducing institutional factors—but in a manner that remains conceptually incomplete and mechanistically underspecified in important ways. Scholars like Jennifer McCoy and Murat Somer have proposed the concept of "pernicious polarization," which adds to measures of affective polarization an

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

assessment of whether political structures incentivize elites to instrumentalize emotional divisions deliberately for electoral gain.[12] This move is significant: it recognizes that the same level of emotional antipathy produces different consequences depending on whether political institutions reward or punish appeals to division. The most polarized and democratically endangered democracies, by this account, are those where affective polarization is high and political leaders have strong incentives to amplify rather than moderate it. This addition of institutional incentives represents a genuine theoretical advance—it explains why some highly affectively polarized societies maintain functioning democratic institutions while others experience democratic backsliding and norm erosion.

Gidron, Adams, and Horne have similarly documented that affective polarization is substantially lower in consensus democracies with proportional representation than in majoritarian systems with single-member districts, despite comparable or even higher levels of inequality, media fragmentation, and demographic diversity.[13] This finding is striking because it isolates the effect of electoral institutions while controlling for socioeconomic and media variables carefully. Democracies with similar structural challenges produce different affective outcomes depending on institutional design choices. This suggests something fundamentally structural: the electoral rules themselves create different incentive environments for political entrepreneurs and politicians. In majoritarian systems where the winner of 50.1 percent of the vote wins the entire district and the opposing 49.9 percent wins nothing, politicians have maximal incentive to mobilize their base intensively, to frame politics in zero-sum identity terms, to build partisan brands rather than coalition-based parties. In proportional systems where 40 percent of the vote yields approximately 40 percent of seats and participation in coalition governments is structurally necessary, politicians face entirely different incentives. They must build bridges to potential coalition partners continuously. They cannot afford to alienate moderate voters or entirely write off minority viewpoints in their calculations. They must maintain relationships across party lines because governance itself requires it. The emotional tenor of politics differs not because citizens are

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

psychologically different across countries, but because the structural incentives facing political entrepreneurs are fundamentally different.

This insight is crucial, yet even here the literature remains frustratingly incomplete and raises more questions than it answers systematically. The mechanisms by which electoral structures shape affective polarization at the mass level are rarely specified in detail or modeled carefully. We know that majoritarian systems correlate with higher affective polarization empirically, but the causal pathways remain undertheorized and the temporal dynamics unclear. Does the electoral rule directly incentivize antagonistic mobilization, or does it work indirectly through other institutions? If so, through what specific mechanisms does electoral incentive translate into citizen affect? Are there feedback loops wherein antagonistic campaign strategies produce affective polarization among voters, which then justifies further antagonistic strategies in the next cycle, creating a self-reinforcing spiral? How do electoral institutions interact with media systems, campaign finance rules, party organization structures, and primary election procedures to produce different affective outcomes? What is the time lag between institutional change and affective change in the mass public?

Most critically, the literature has not yet grappled adequately with what we might call the "design architecture" problem: the fact that electoral systems are not natural phenomena but deliberately constructed institutional arrangements, and that these arrangements embed assumptions about how politics should be conducted, what voters care about, what forms of mobilization are permissible, and what constitutes legitimate political conflict. An electoral system is not merely a neutral technical rule about how votes translate to seats. It is an entire apparatus that affords certain political behaviors and constrains others, that makes certain strategies more advantageous than others, that shapes the very ecology within which political actors operate and respond to incentives. Understanding institutions as designed systems—as objects that could be designed

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

differently and deliberately—opens entirely new conceptual possibilities for thinking about polarization and institutional reform.

The Epistemic Stratification Gap

Running parallel to the institutional mechanism problem is a second, largely separate and understudied lacuna in the polarization literature: the near-complete absence of serious engagement with what we might term epistemic stratification, or the structural fragmentation of knowledge systems that prevents the formation of a shared epistemic space for democratic reasoning and deliberation.

The polarization literature has devoted considerable attention to misinformation, disinformation, and information bubbles in recent years. Media scholars have documented the rise of partisan news sources, the decline of trusted local journalism, the algorithmic curation of social media feeds toward emotionally engaging content.[14] These explanations correctly identify a real problem in contemporary media systems: citizens increasingly consume news filtered through partisan lenses, fact-checks often fail to convince those who disagree, and false information spreads rapidly through social networks. The proposed solution follows logically: improve information quality, increase media literacy, reduce algorithmic amplification of false claims, rebuild local journalism as a trusted source of record. Numerous interventions proceed from this diagnosis: fact-checking programs, media literacy initiatives, platform algorithm modifications.

Yet this framework mistakes the fundamental problem at its core. The actual issue is not primarily that citizens lack accurate information—though that is sometimes true—but rather that democratic societies have become organized around fundamentally incompatible standards for what counts as knowledge, which expertise is legitimate, and how collective reasoning should proceed systematically. Expert knowledge, produced through academic peer review and published in specialized journals, operates according to one set of evidentiary standards and addresses one

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

epistemic community of specialists. Political knowledge, produced through legislative deliberation and media coverage of political events, operates according to different standards—where narratives, electoral consequences, and institutional constraints shape what becomes politically salient and actionable. Citizen knowledge, constructed through personal experience, family networks, and the media consumed in daily life, follows yet a third logic entirely. These three epistemic strata do not translate cleanly into one another without loss or distortion. A finding published in a peer-reviewed journal does not automatically become "knowledge" in the political domain; it must be translated, simplified, selectively emphasized or de-emphasized depending on political consequences. Political knowledge does not filter cleanly down to citizens; it passes through media institutions that editorialize, frame, and interpret according to their own institutional logics and business incentives.

Moreover, and this is the critical point rarely made explicit in polarization literature, these stratified epistemic systems are themselves products of institutional design—deliberate choices made by people building institutions to accomplish particular goals. The peer review system that certifies expert knowledge was designed with particular purposes in mind—establishing quality control, enabling cumulative knowledge production, creating credential-based authority structures. Legislative deliberation operates according to procedures and norms designed to enable coalition-building and compromise. Media institutions evolved with particular economic models and professional norms that shaped what they cover and how they cover it. And crucially, when these institutional designs change—when academic journals fragment into specialized silos, when cable news replaces local journalism as a primary news source, when social media replaces editorial gatekeeping—the epistemic landscape transforms fundamentally. Not because citizens become stupider or more conspiratorial, but because the institutional infrastructure enabling shared reasoning has been dismantled or transformed.

The polarization literature has not adequately theorized this problem, instead treating it as a failure

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

of individual cognition or media business models alone. Scholars recognize that different groups "live in different realities" or consume "different information," and they correctly note that misinformation is a problem. But they rarely ask the deeper question: what institutional conditions would need to be in place for genuinely shared epistemic space to exist? What would have to change about how we produce, validate, and disseminate knowledge such that citizens operating from different social positions could nonetheless engage in collective reasoning about shared problems? How would institutions need to be redesigned such that expert knowledge, political deliberation, and citizen knowledge could translate into one another intelligibly? This gap in the literature represents not merely an oversight but a fundamental misunderstanding of what polarization is. Polarization is not primarily a failure of individual rationality or even of information quality. It is a structural consequence of epistemic architecture that prevents different knowledge communities from reasoning together.

The Individual-Collective Disconnect and Feedback Loop Dynamics

Perhaps the most significant yet underexplored puzzle in polarization research concerns the disconnect between individual-level attitude change and collective-level democratic outcomes and system behavior. The empirical fact is now well-established through numerous rigorous studies: you can reduce an individual's hostility toward the opposing party, increase their accurate perception of where their policy views actually overlap, even facilitate genuine friendly contact across party lines—and none of this reliably translates into reduced polarization at the collective level or improved democratic functioning. The most recent experimental evidence from the Strengthening Democracy Challenge, which tested hundreds of interventions on tens of thousands of participants systematically, found that reducing affective polarization had virtually no effect on willingness to support democratic norms or reject antidemocratic candidates.[15] The promising laboratory findings simply did not scale to the real political world.

The existing literature explains this primarily through reference to social context and

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

environmental reinforcement: yes, the dinner guests become friendlier during the moderated dialogue, but they return to their religious congregations, workplaces, and media environments that pull them back toward polarized orientations. The affective gains decay as people are reimmersed in environments that reward hostility and punish moderation systematically. This explanation is not wrong, but it is incomplete and misses the systemic level. It remains trapped within individual psychology—the atomized person, surrounded by contextual forces, unable to sustain changed attitudes against environmental pressure. It assumes that the problem is individual-level but context-shaped, rather than genuinely systemic in nature.

What the literature has failed to adequately theorize is that collective polarization is not the sum of individual affective states, but rather an emergent property of system-level dynamics and feedback mechanisms. This is not a metaphor but a literal claim grounded in systems dynamics theory: polarization at the collective level is produced by reinforcing feedback loops, institutional incentive structures, and equilibria that persist even when individual preferences shift. Consider the electoral feedback loop in its full complexity and temporal unfolding. Majoritarian electoral systems incentivize political parties to mobilize their base intensively, to frame politics in identity terms, to emphasize what divides rather than unites. These campaign strategies are not chosen because politicians are innately divisive or malevolent, but because they are rational given the structural incentives they face. Voters, exposed to antagonistic campaign messaging from multiple cycles, gradually sort themselves into identity-based camps where partisan identity becomes fused with other identities. Once sorted into camps, they consume media that confirms their camp identity and reinforces hostility toward the opposing camp. Their affective polarization increases. This increased affective polarization justifies further antagonistic campaigning ("you see how much the other side hates us? They're attacking our way of life"). This fuels further sorting and affective division. The system locks into a reinforcing cycle that persists at the collective level. Any individual who reduces their own hostility or becomes aware of policy overlap faces substantial social pressure from their camp—their family, their church, their neighborhood, their media diet—pulling them back toward

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

antagonism. The individual change does not propagate upward to shift the system because the system-level incentives remain unchanged and structurally different.

This dynamic is well understood in systems dynamics theory and has been extensively modeled in the literature on complex systems and feedback loops.[16] Yet the polarization literature has largely failed to engage with this theoretical tradition systematically. Most polarization research operates in the idiom of individual psychology—attitudes, beliefs, preferences, behaviors, emotions. The systemic properties of feedback loops, equilibria, and emergent dynamics remain largely invisible or treated as background context rather than as the primary object of analysis. Consequently, the field has developed a systematic bias toward individual-level solutions (change attitudes, correct misperceptions, bring people together for dialogue) that are structurally incapable of addressing system-level problems (electoral incentives that reward antagonism, institutional architectures that prevent shared reasoning, feedback mechanisms that lock polarization in place). Overcoming this bias requires genuine theoretical reorientation.

The Temporal and Spatial Lacunae

Two additional gaps deserve mention, though they are less central than the institutional and epistemic problems: the almost complete absence of temporal analysis and the near-total neglect of spatial and material dimensions of polarization.

On the temporal dimension: polarization research is predominantly cross-sectional or produces long-term trend data. We have excellent snapshots of polarization at particular moments, comprehensive trend data on how polarization has evolved over decades, and experimental evidence on the short-term effects of interventions. What we almost entirely lack is theorization of the dynamics of polarization over time—the actual mechanisms through which polarization accelerates, plateaus, or decelerates. When and how does polarization harden into something that becomes path-dependent and resistant to reversal? What conditions allow depolarization or

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

democratization after periods of polarization? Are there inflection points beyond which polarization becomes self-sustaining and extremely difficult to reverse? The question is not merely descriptive but mechanistic: what temporal dynamics characterize polarized systems as distinct from depolarizing systems? The literature provides almost no guidance. We know that affective polarization emerged gradually from the 1980s onward, that it accelerated in some periods, that it shows asymmetric trajectories across parties. But we lack a theoretical framework for understanding polarization as a temporal process with its own dynamics.

Similarly, polarization research operates in a largely aspatial and dematerialized register. We discuss polarization as if it were a property of abstract political attitudes, divorced from the places where people actually live, work, and interact. Yet polarization is deeply geographical. It is structured by whether one lives in dense urban cores or sparse rural areas, by economic geography and regional inequality, by the physical spaces where people encounter one another or deliberately avoid one another.[17] It is shaped by the material infrastructure of political participation—whether there are accessible polling places, whether local newspapers still exist, whether there are public spaces where people from different backgrounds encounter one another. The material and spatial dimensions of polarization remain almost entirely absent from scholarly analysis, replaced by a disembodied focus on attitudes, information, and psychological processes.

The Design Methodology Gap: Toward a Framework of Institutional Specifications

The preceding gaps—the underspecification of institutional mechanisms, the neglect of epistemic architecture, the failure to theorize feedback loops, the temporal and spatial lacunae—point toward a fundamental reframing of the polarization problem itself. What has been missing is a conceptual framework that treats polarization not as a psychological or informational problem amenable to attitude change through intervention, but as a problem of institutional design—understood in the rigorous sense employed in design theory and systems thinking by design scholars.

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

Design methodology, developed primarily in architecture, industrial design, and human-centered design disciplines, offers conceptual tools and analytical approaches largely absent from political science literature and political analysis. Central is the concept of affordances: the possibilities for action that an object, system, or environment makes available to an actor.[18] An electoral system affords certain forms of political behavior. A majoritarian system affords antagonistic zero-sum mobilization; it constrains coalition-building and incentivizes identity-based party formation. A proportional system affords different behaviors—coalition-building becomes necessary, moderate voices matter, antagonism carries electoral costs. These are not psychological or individual-level properties; they are structural properties of the institutional landscape. Actors respond rationally to the affordances available to them. If the system affords antagonism and punishes moderation, rational actors become antagonistic. Change the system, and you change the range of rational behaviors available and rational strategies for politicians to pursue.

Similarly, design methodology emphasizes the distinction between treating a phenomenon as a problem to be solved (implying a solution exists if we are clever enough) versus a wicked problem that cannot be solved but only managed through ongoing iteration and redesign.[19] Polarization as currently framed—as a problem of individual attitudes, media effects, or information quality—is treated as a solvable problem amenable to a definitive solution. Get better information to people, change media incentives, modify voter attitudes, and polarization diminishes. The design methodology perspective suggests this is a fundamental category error. Polarization emerges from multiple coupled systems (electoral incentives, media structures, knowledge production, campaign finance, party organization, voter sorting) that interact in complex ways. No single-factor intervention can address it because the system compensates. Change media content without changing electoral incentives, and politicians still have reason to polarize. Change electoral incentives without addressing epistemic stratification, and groups still cannot reason together. The problem is wicked precisely because it is systemic. The question is not whether perfect depolarization is achievable, but what institutional specifications would create conditions where

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

different affective and epistemic outcomes become possible.

Finally, design methodology emphasizes the importance of understanding what becomes possible under different design configurations. Rather than asking "what is the problem with these people?" we ask "what does this system afford?" and "what would become possible under a different design?" This is not utopianism but rigorous analysis grounded in design theory. The question is not whether perfect democracy is achievable, but what institutional specifications would create conditions where political actors have incentives toward moderation, where knowledge communities can reason together, where citizens experience politics as something other than existential conflict.

Synthesis: Toward an Integrated Framework

What emerges from mapping these research gaps comprehensively is not a minor omission but a fundamental reconceptualization of polarization as an institutional design problem at its core. The argument, stated directly and clearly, is this: political polarization in liberal democracies is not primarily a psychological phenomenon, an informational failure, or an elite conspiracy, though all three factors play roles in the overall system dynamics. It is fundamentally a structural outcome of specific institutional design choices—namely, winner-take-all electoral systems that incentivize antagonistic mobilization; epistemically stratified knowledge systems that prevent collective reasoning; and affective infrastructures (campaign finance, media attention allocation, party organization, primary election procedures) that reward in-group solidarity and out-group hostility. These systems interact through reinforcing feedback loops that lock polarization in place structurally. Because polarization is produced by coupled systems, not individual attitudes, individual-level interventions cannot address it fundamentally. The system simply compensates by redirecting the individual's affective orientation back toward the equilibrium state. The problem is not individual irrationality but structural incentives embedded in institutional design.

This reframing opens an entirely different set of research questions and policy possibilities

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

previously unimaginable. Rather than asking "how do we change individual attitudes?" we ask "what institutional specifications would afford different affective and epistemic possibilities?" Rather than treating polarization as pathological emotion to be cured, we ask "what democratic values are we willing to trade for different institutional designs?" Rather than seeking the one intervention that will fix everything, we recognize that managing polarization in democratic systems requires attention to multiple coupled systems: electoral rules that create incentives for coalition-building rather than antagonism, epistemic institutions that enable translation between different knowledge communities, media systems that rebuild local journalism as a trusted source, party organizations that depend on coalition-building rather than base mobilization. The goal is not depolarization as an absolute state, but rather the creation of institutional conditions within which democratic governance becomes possible despite differences and disagreement.

The literature on polarization has made genuine progress over recent decades. The first-generation recognition that polarization is a measurable phenomenon that has increased over time, the second-generation insight that polarization is affective rather than merely ideological, that it involves identity-based reasoning, that it is shaped by elite behavior and media institutions—these represent real advances. Yet the field remains trapped within frameworks that, despite their sophistication, cannot explain the core puzzles: why institutional variation produces such radically different outcomes, why individual interventions fail to scale from laboratory to society, why we see such divergent temporal trajectories of political violence even where affective polarization is similar, and why the same affective polarization level is compatible with both functioning and failing democracies depending on institutional context.

Addressing this gap requires drawing on theoretical traditions rarely engaged by polarization scholars: comparative institutional analysis that specifies mechanisms rather than merely correlating institutions with outcomes; design methodology that treats institutions as artifacts produced through deliberate choices and amenable to redesign; systems dynamics that models

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

feedback loops and emergent properties; and epistemic analysis that treats knowledge production and validation as institutionally structured processes. The resulting framework would not solve polarization—wicked problems are not solved but managed—but it would provide the conceptual apparatus to ask better questions and propose more plausible approaches to democratic governance in polarized contexts. This represents not a minor amendment to existing scholarship but a fundamental reorientation of how we think about polarization and democratic possibility. Future research grounded in this framework could generate genuinely novel insights into democratic possibility.

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

The path forward requires institutional experimentation and creative redesign that extends beyond conventional political science analysis. This does not mean abandoning the gains of contemporary scholarship—the insights about affective polarization, elite incentives, and media effects remain valid and important. Rather, it means integrating these insights into a larger framework that explains how individual psychology, institutional design, media systems, and epistemic architecture interact dynamically to produce polarized democracies. Such integration opens new possibilities for both research and practical democratic innovation. Scholars can move beyond asking whether interventions work in isolation toward asking how institutional changes might alter the incentive landscape within which both elites and ordinary citizens operate. Policymakers might move beyond seeking the single policy fix toward thinking seriously about coupled institutional redesign that addresses multiple systems simultaneously. Democratic theorists might move beyond treating polarization as a pathology to be cured toward understanding it as an ongoing management challenge requiring continuous institutional attention and democratic adaptation. This reframing does not offer false hope of total depolarization, but it does suggest more plausible paths toward functional democratic governance within persistently plural and disagreeing societies. The intellectual work of developing this integrative framework and the practical work of testing alternative institutional configurations remains largely ahead. Yet mapping the conceptual gaps as this review does is an essential first step toward genuine scholarly and democratic possibility.

Works Cited

Abramowitz, Alan I., and Kyle L. Saunders. "Is Polarization a Myth?" Journal of Politics, vol. 70, no. 2, Apr. 2008, pp. 542–555.

Bishop, Bill. The Big Sort: Why the Clustering of Like-Minded America Is Tearing Us Apart. Houghton Mifflin Harcourt, 2008.

Boxell, Levi, Matthew Gentzkow, and Jesse M. Shapiro. "Cross-Country Trends in Affective Polarization." National Bureau of Economic Research Working Paper, no. 26669, 2020.

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

Broockman, David, Joshua Kalla, and Sean Westwood. "Does Affective Polarization Undermine Democratic Norms or Accountability? Maybe Not." American Journal of Political Science, online first, 2023.

Fiorina, Morris P., Samuel A. Abrams, and Jeremy C. Pope. Culture War? The Myth of a Polarized America. Pearson Longman, 2010.

---. "Polarization in the American Public: Misconceptions and Misreadings." Journal of Politics, vol. 70, no. 2, 2008, pp. 556–560.

Gidron, Noam, James Adams, and Will Horne. American Affective Polarization in Comparative Perspective. Cambridge University Press, 2020.

Gimpel, James G., and Iris S. Hui. "Seeking Politically Compatible Neighbors? The Role of Neighborhood Partisan Composition in Residential Sorting." Political Geography, vol. 48, Sept. 2015, pp. 130–142.

Hall, Andrew B. Who Wants to Run? How the Devaluing of Political Office Drives Polarization. University of Chicago Press, 2019.

Iyengar, Shanto, Gaurav Sood, and Yphtach Lelkes. "Affect, Not Ideology: A Social Identity Perspective on Polarization." Public Opinion Quarterly, vol. 76, no. 3, 2012, pp. 405–431.

Kleinfeld, Rachel. Polarization, Democracy, and Political Violence in the United States: What the Research Says. Carnegie Endowment for International Peace, 2023.

Kubin, Emily, and Christian Von Sikorski. "The Role of (Social) Media in Political Polarization: A Systematic Review." Annals of the International Communication Association, vol. 45, no. 3, 2021, pp. 188–206.

Lee, Frances E. Beyond Ideology: Politics, Principles, and Partisanship in the U.S. Senate. University of Chicago Press, 2009.

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

Levendusky, Matthew. "Partisan Media Exposure and Attitudes Toward the Opposition." Political Communication, vol. 30, no. 4, 2013, pp. 565–582.

Martin, Gregory J., and Ali Yurukoglu. "Bias in Cable News: Persuasion and Polarization." American Economic Review, vol. 107, no. 9, 2017, pp. 2,565–2,599.

Mason, Lilliana. Uncivil Agreement: How Politics Became Our Identity. University of Chicago Press, 2018.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. Polarized America: The Dance of Ideology and Unequal Riches. MIT Press, 2006.

McCoy, Jennifer, Tahmina Rahman, and Murat Somer. "Polarization and the Global Crisis of Democracy." American Behavioral Scientist, vol. 62, no. 1, 2018, pp. 3–17.

McCoy, Jennifer, and Murat Somer. "Toward a Theory of Pernicious Polarization and How It Harms Democracies." The ANNALS of the American Academy of Political and Social Science, vol. 681, no. 1, 2019, pp. 234–271.

Meadows, Donella H. Thinking in Systems: A Primer. Chelsea Green, 2008.

Norman, Donald. The Design of Everyday Things. MIT Press, 2013.

Nyhan, Brendan, et al. "Like-Minded Sources on Facebook Are Prevalent but Not Polarizing." Nature, vol. 620, Aug. 2023, pp. 137–144.

Paluck, Elizabeth L., et al. "Prejudice Reduction: Progress and Challenges." Annual Review of Psychology, vol. 72, 2021, pp. 563–590.

Poole, Keith T., and Howard Rosenthal. Congress: A Political-Economic History of Roll-Call Voting. Oxford University Press, 1997.

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved

Rittel, Horst W. J., and Melvin M. Webber. "Dilemmas in a General Theory of Planning." Policy Sciences, vol. 4, no. 2, 1973, pp. 155–169.

Stewart, Alexander J., Nolan McCarty, and Joanna J. Bryson. "Polarization Under Rising Inequality and Economic Decline." Science Advances, vol. 6, no. 50, Dec. 2020, article eabd4201.

Suchman, Lucy. Human-Machine Reconfigurations. Cambridge University Press, 2007.

Tajfel, Henri, and John Turner. "An Integrative Theory of Intergroup Conflict." The Social Psychology of Intergroup Relations, edited by William G. Austin and Stephen Worchel, Brooks/Cole, 1979.

Voelkel, Jan G., et al. "Interventions Reducing Affective Polarization Do Not Necessarily Improve Anti-Democratic Attitudes." Nature Human Behaviour, vol. 7, 2023, pp. 1–10.

---. "Megastudy Identifying Effective Interventions to Strengthen Americans' Democratic Attitudes." Proceedings of the National Academy of Sciences, vol. 119, no. 20, May 2022, article e2120974119.

https://doi.org/10.33774/apsa-2026-xhl41 **ORCID:** https://orcid.org/0009-0005-2942-0839 Content not peer-reviewed by APSA. **License:** All Rights Reserved