(REVIEW ARTICLE)

Disinformation as a driver of political polarization: A strategic framework for rebuilding civic trust in the U.S

Mohammed Hafiz Nabila [1], Rohany Abdul Shaibu [2], Glory Edinam Afeti [3, *] and Esinu Aku Adza [4]

[1] Department of Communication Studies and Storytelling, East Tennessee State University, USA.
[2] Department of Accounting, University of Ghana.
[3] Department of Development Policy, Ghana Institute of Management and Public Administration, Ghana.
[4] Institute of Statistical, Social and Economic Research, University of Ghana.

Abstract

This study investigates the strategic role of disinformation in exacerbating political polarization and eroding civic trust in the United States. In the context of increasing ideological fragmentation and declining institutional legitimacy, the paper aims to develop a comprehensive framework for understanding how disinformation, both foreign and domestic, functions as a systematic tool of democratic destabilization. The study synthesizes existing theoretical and empirical literature from political science, media studies and computational social science. This paper draws on secondary data, content analyses and findings from peer-reviewed studies to explore five core dimensions: the theoretical foundations of political polarization in the U.S., the structural design of disinformation ecosystems, cognitive mechanisms underlying belief formation, digital media architectures and algorithmic amplification and institutional responses aimed at democratic resilience. The findings reveal that algorithmically driven digital platforms disproportionately amplify partisan and emotionally charged content; however, cognitive biases such as confirmation bias and motivated reasoning make individuals more susceptible to politically aligned disinformation. The study also highlights how foreign actors exploit domestic vulnerabilities through computational propaganda and how domestic political actors weaponize disinformation to manipulate public opinion and suppress dissent. Institutional responses such as fact-checking and media literacy programs, though valuable, remain insufficient in the absence of systemic platform reform and coordinated governance mechanisms. The paper, therefore, concludes that disinformation operates not merely as a crisis of information accuracy but as an epistemic threat to democratic deliberation. Addressing this challenge requires an integrated strategy that combines policy, technology, education, and civic renewal.

Keywords: Disinformation; Political Polarization; Civic Trust; Democratic Resilience

1. Introduction

The contemporary American political landscape is marked by unprecedented levels of partisan division, ideological separation and epistemic fragmentation that threaten the foundational principles of democratic governance (Shelley, 2022). At the center of this democratic crisis lies a phenomenon of significant scholarly and practical importance: the systematic spread of disinformation and its role as both a catalyst and amplifier of political polarization. This paper explores the complex dialectical relationship between disinformation campaigns, political polarization and the erosion of civic trust, however, proposing a comprehensive strategic framework for democratic renewal in the twenty-first century. According to Dahlgren (2018), the United States faces what scholars have termed an "epistemic crisis, a fundamental breakdown in shared methods of distinguishing truth from falsehood that undermines the deliberative foundations of a democratic society. Conceptually, this crisis manifests most acutely in the realm of political discourse,

* Corresponding author: Glory Edinam Afeti

where competing information ecosystems have emerged, each with distinct epistemological frameworks, factual claims and interpretive schemas. The proliferation of digital communication technologies, coupled with the strategic deployment of disinformation by both domestic and foreign actors, has accelerated the fragmentation of the American public sphere into what Marino et al (2024) characterize as "echo chambers" and "information cocoons".

Recent empirical research shows that political polarization in the United States has reached levels not seen since the Civil War era (Kleinfeld, 2023). Kleinfeld (2023) underscored that with affective polarization, emotional reactions toward opposing partisans exhibit particularly dramatic increases. Besides, trust in democratic institutions has sharply declined, with surveys indicating that substantial majorities of Americans express skepticism about the integrity of electoral processes, the credibility of mainstream media and the legitimacy of governmental authority (Citrin & Stoker, 2018). This convergence of polarization and institutional distrust creates what political scientists term a "crisis of democratic legitimacy," which poses existential challenges to the American constitutional order. The challenge of disinformation-driven polarization represents one of the most urgent threats to American democracy in the twenty-first century. Through viewing disinformation not merely as a collection of false statements but as a systematic strategy for democratic destabilization, this research provides new insights into both the nature of contemporary threats to democracy and the institutional innovations needed to address them. The urgency of this investigation is highlighted by ongoing events that reveal the real-world consequences of an epistemic crisis, from contested elections to public health misinformation to climate change denial. The researchers believe that denveloping effective responses to disinformation whilest preserving democratic values demands comprehensive, theoretically grounded and empirically rigorous analysis. The stakes could not be higher: the future of American democracy may well depend on our collective ability to rebuild shared epistemic foundations for democratic deliberation and governance.

2. Literature review

The scholarly examination of disinformation's impact on political polarization requires synthesis across multiple disciplinary domains, including political science research on democratic legitimacy and institutional trust, communication studies on media effects and information processing, and emerging interdisciplinary work on computational propaganda and digital authoritarianism. Whereas substantial bodies of literature exist on political polarization as a standalone phenomenon and disinformation as a discrete communication strategy, relatively few studies have systematically theorized their dynamic interaction or developed comprehensive frameworks for understanding their combined effects on democratic governance. This literature review, therefore, integrates findings from disparate research traditions to construct a theoretical foundation for examining disinformation-driven polarization as a distinct threat to civic epistemology and democratic resilience.

2.1. Theoretical Foundations of Political Polarization in the USA: Elite, Mass, and Affective Dimensions

The theoretical understanding of American political polarization has evolved through three distinct but interconnected analytical frameworks that brighten different dimensions of partisan division within democratic governance (Baliñas Pérez, 2024). Elite polarization theory, pioneered by Poole and Rosenthal's (1984, 2007) NOMINATE scaling methodology, conceptualizes polarization as the ideological divergence between political elites as measured through congressional voting behaviour, which demonstrates that Republican and Democratic legislators have moved to increasingly distant ideological positions since the 1970s. This institutional approach, extended by McCarty et al (2006) in "Polarized America," emphasizes how primary election dynamics, partisan gerrymandering and congressional institutional changes create incentive structures that reward ideological purity and punish cross-party cooperation. This generates what they term "partisan-ideological polarization" that becomes self-reinforcing over time through strategic calculations of political actors operating within specific electoral and legislative contexts.

Alternatively, mass polarization theory examines whether elite divisions translate into corresponding ideological polarization among ordinary citizens, generating intense scholarly debate between competing theoretical perspectives on public opinion formation and democratic representation (Robison & Mullinix, 2016). The "polarized public" perspective, exemplified by Bolinas Pérez (2024), argues that politically engaged citizens have indeed adopted more extreme positions and exhibit increasing bimodality in ideological self-identification. Even though the "sorted public" framework developed by Baldassarri & Page (2021) contends that apparent polarization reflects citizens aligning their partisan identities with pre-existing policy preferences rather than genuine attitudinal extremism. Conceptually, this theoretical distinction between polarization and sorting carries profound implications for understanding democratic responsiveness, as genuine mass polarization would suggest that elite polarization accurately represents constituent preferences, while mere sorting would indicate a failure of democratic representation that contributes to governance dysfunction.

Affective polarization theory represents the most recent theoretical development, which focuses on emotional reactions and social attitudes between partisan groups rather than policy disagreements or ideological distances (Webster & Abramowitz, 2017). This is drawn from social identity theory to understand partisanship as a form of social identity that generates in-group favouritism and out-group hostility independent of policy considerations. King Zette (2021) "Uncivil Agreement" provides the theoretical foundation for this approach, which argues that American politics exhibits "social polarization" characterized by intense partisan animosity. Despite relatively moderate policy preferences, as partisan identities become "stacked" with other social identities, including race, religion and geography, to create reinforcing group boundaries. This framework is supported by experimental evidence from Blasco ich, (2017), which demonstrates that partisan identities influence behaviour across non-political domains, conceptualizes contemporary American polarization as fundamentally driven by social identity processes that heighten intergroup competition and emotional responses toward opposing partisans, providing essential theoretical groundwork for understanding how disinformation might interact with existing polarization dynamics through elite strategic resources, mass opinion formation processes, and social identity activation mechanisms.
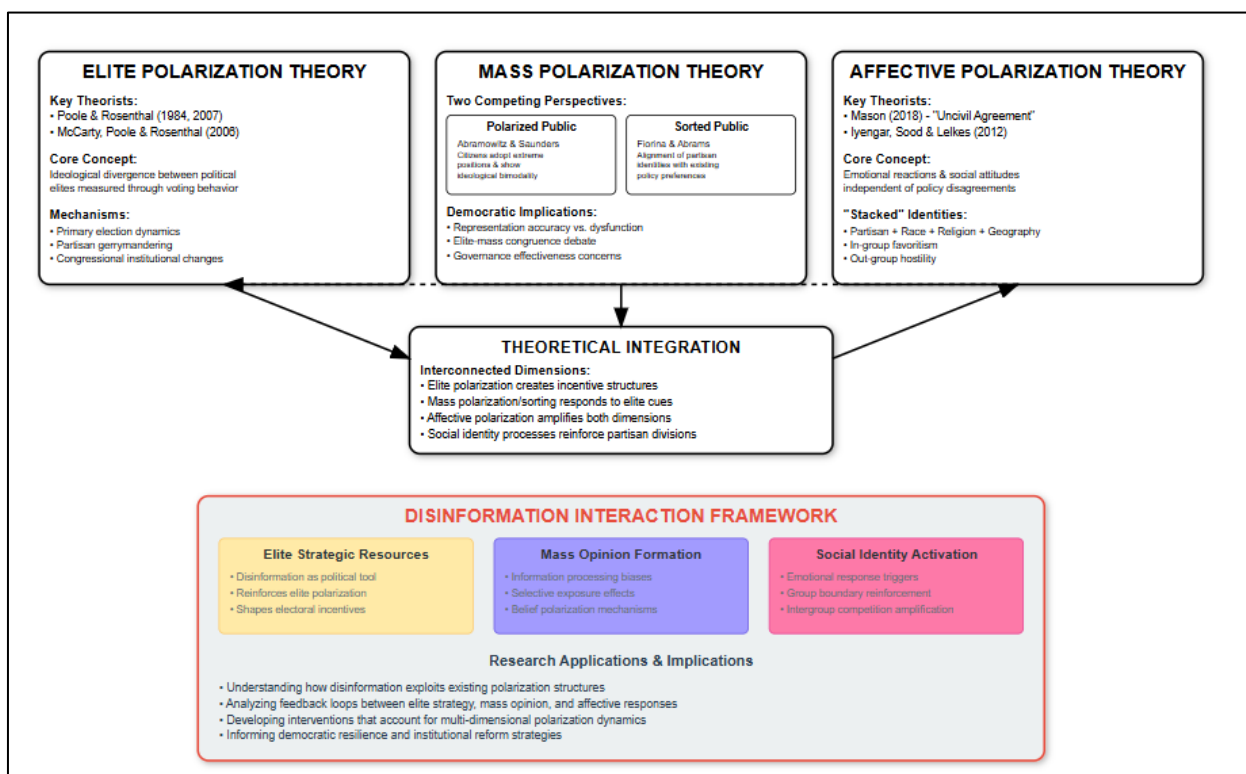

Figure 1 Theoretical Foundations of Political Polarization in the USA

The chart above illustrates the tripartite theoretical framework of American political polarization, which encompasses elite polarization theory (focused on institutional dynamics and strategic positioning), mass polarization theory (examining the scholarly debate between "polarized public" versus "sorted public" perspectives) and affective polarization theory (emphasizing social identity processes and emotional responses toward opposing partisans). The framework demonstrates theoretical integration through interconnected mechanisms where elite polarization creates incentive structures, mass polarization responds to elite cues through sorting processes and affective polarization amplifies both dimensions via social identity activation that reinforces partisan divisions across multiple social domains. The lower section reveals how disinformation strategically exploits each polarization dimension through distinct pathways, which provide elite actors with strategic resources for political manipulation, thus influencing mass opinion formation through selective exposure and belief polarization mechanisms and activating social identity responses that heighten intergroup competition and emotional hostility. This comprehensive framework enables systematic analysis of how disinformation both capitalizes on existing polarization structures and generates feedback loops that intensify democratic dysfunction across elite, mass, and affective dimensions of American political life.

2.2. Disinformation Ecosystems and Information Warfare: Conceptual Frameworks and Empirical Evidence from the USA

The conceptualization of disinformation as systematic information warfare rather than isolated instances of false information represents a fundamental theoretical shift in understanding how deliberate falsehoods function within democratic societies. Livingston & Bahador (2025), in "Network Propaganda," provide the foundational framework for understanding American disinformation ecosystems as asymmetric network structures, where right-wing media operates as an insular "propaganda feedback loop" that amplifies false narratives. However, left-wing and centrist media maintain greater connections to professional journalism norms and fact-checking institutions. The analysis of media linking patterns, social media sharing behaviours and content analysis during the 2016 election demonstrates that disinformation propagates through distinct structural pathways characterized by high clustering coefficients, limited bridge connections to mainstream media, and algorithmic amplification mechanisms that create self-reinforcing echo chambers (Livingston & Bahador, 2025). This ecosystem approach emphasizes that the effectiveness of disinformation depends on the credibility of individual false claims and the structural properties of information networks that determine message reach, repetition frequency and source credibility within specific audience segments.

Empirical evidence from the American disinformation landscape reveals sophisticated operational frameworks that integrate computational propaganda techniques, strategic narrative construction, and psychological manipulation tactics to achieve specific political objectives. Marsden et al. (2020) document the Russian Internet Research Agency's systematic approach during the 2016 election, which combined demographic microtargeting, cultural grievance amplification and coordinated inauthentic behaviour to exploit existing social divisions and suppress voter turnout among key Democratic constituencies. Their analysis reveals that effective disinformation campaigns operate through what they term "informational gerrymandering"; the strategic manipulation of information flow to create artificial consensus within targeted communities, whereas maintaining plausible deniability about coordinated manipulation. Vosoughi et al. (2018) conducted a comprehensive analysis of Twitter diffusion patterns that provides crucial empirical support for disinformation effectiveness, which demonstrates that false information spreads six times faster than accurate information and reaches more people through network effects that prioritize novelty and emotional arousal over factual accuracy. This suggests that disinformation possesses inherent structural advantages within digital communication environments optimized for engagement rather than truth.

According to Whyte (2020), the strategic dimensions of information warfare in the American context encompass both foreign interference operations and domestic disinformation campaigns that exploit democratic institutions' vulnerabilities to deliberate manipulation. Ferrara (2020) establishes the theoretical framework for understanding "computational propaganda" as the algorithmic manipulation of public opinion through automated content generation, bot networks and platform gaming techniques that amplify minority viewpoints to create false impressions of grassroots support. Domestic disinformation operations, documented extensively by Ferrara (2020) in "Cyberwar," demonstrate how political actors weaponize information asymmetries to undermine electoral integrity, scientific consensus and institutional credibility through coordinated campaigns that blend legitimate political communication with deliberate falsehoods designed to create epistemic uncertainty and democratic dysfunction. This strategic approach to disinformation transcends traditional propaganda models by targeting the epistemological foundations of democratic deliberation itself, which creates what scholars' term "truth decay" (Beaumont & Beerbohm, 2024). The distinction between factual claims and opinion becomes increasingly blurred, thereby undermining citizens' capacity for informed political participation and democratic accountability mechanisms essential for legitimate governance.

2.3. Cognitive Mechanisms of Belief Formation: Motivated Reasoning, Confirmation Bias and Epistemic Closure

According to Cushman (2020), People form and maintain beliefs through cognitive mechanisms that often deviate from rational information processing. Motivated reasoning occurs when individuals direct their mental resources toward conclusions that support their existing preferences and identities (Wisniewski, 2022). People apply more rigorous scrutiny to information that challenges their beliefs, but use lenient standards for confirming evidence. This creates systematic bias in how beliefs are updated. Individuals become skilled at generating justifications for their preferred positions while identifying flaws in contradictory evidence. Inferences from the literature revealed that brain imaging studies show that motivated reasoning activates reward centres when people successfully defend their existing beliefs. This suggests that maintaining cognitive consistency provides psychological rewards that reinforce biased thinking.

Confirmation bias operates through various cognitive pathways that influence how people search for, interpret and remember information. This bias appears at multiple stages of thinking. First, individuals seek information sources that align with their existing beliefs. Second, they interpret ambiguous evidence as supporting their prior views. Third, they recall confirming instances more easily than disconfirming ones (Pilditch, 2017). Confirmation bias becomes stronger under cognitive load and time pressure, as people rely on mental shortcuts. Modern research indicates that confirmation

bias extends beyond individual limitations. Social and technological factors amplify it, including algorithms that create "filter bubbles" and partisan media that provide ideologically consistent narratives (Akola, 2025; Papp, 2024).

According to Kruglanski et al (2018), epistemic closure represents the most extreme form of biased belief formation. It involves creating self-contained belief systems that resist external challenges. This occurs when people develop comprehensive worldviews that dismiss contradictory evidence. Epistemic closure comprises several components. People establish trusted authorities as exclusive sources of legitimate knowledge. They develop beliefs about the unreliability of mainstream sources. They create alternative explanatory frameworks that address apparent contradictions. The digital age has enabled unprecedented levels of epistemic closure through fragmented information ecosystems and parallel institutional structures. Understanding these three mechanisms is essential for developing interventions to promote accurate belief formation. Each mechanism functions through different cognitive and social pathways that require distinct approaches.

Table 1 Empirical Evidence: Key US Academic Sources

Source	Institution/Publisher	Key Findings	Relevance
Nickerson, R.S. (1998). "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises"	*Review of General Psychology*, American Psychological Association	Comprehensive review demonstrating confirmation bias across multiple domains of human reasoning	Foundational empirical evidence for confirmation bias mechanisms
Westen, D. et al. (2006). "Neural Bases of Motivated Reasoning"	*Journal of Cognitive Neuroscience*, MIT Press	Neuroimaging evidence shows reward activation during motivated reasoning	Biological basis for motivated reasoning processes
Klayman, J. & Ha, Y.W. (1987). "Confirmation, Disconfirmation, and Information"	*Psychological Review*, American Psychological Association	Experimental evidence of biased hypothesis testing strategies	Core research on information processing biases
Kunda, Z. (1990). "The Case for Motivated Reasoning"	*Psychological Bulletin*, American Psychological Association	Theoretical framework for motivated cognition processes	Foundational theory on motivated reasoning
Sanchez, J. (2009). "Epistemic Closure and the Conservative Movement"	*Cato Institute*, Washington DC	Analysis of epistemic closure in political movements and information ecosystems	Application of epistemic closure to contemporary politics

The table above demonstrates the robust empirical foundation underlying the three cognitive mechanisms of belief formation discussed in this section. The convergence of evidence from experimental psychology (Klayman & Ha, Nickerson), neuroscience (Westen et al.) and political analysis (Sanchez) illustrates how motivated reasoning, confirmation bias, and epistemic closure operate across different levels of analysis, from individual neural processes to broader social and political phenomena. The scholarly consensus represented in these publications, spanning over two decades of research from premier American academic institutions, validates the theoretical framework, which highlights the practical implications for understanding contemporary challenges in democratic discourse and public reasoning.

2.4. Digital Media Architectures and Algorithmic Amplification of Partisan Content in the USA

The architecture of digital media platforms fundamentally shapes information consumption patterns through algorithmic curation systems designed to maximize user engagement (Lee, 2024). Social media platforms employ machine learning algorithms that analyze user behavior data to predict and deliver content most likely to generate clicks, shares and extended viewing time (Phillips et al. 2017). These engagement-driven algorithms create feedback loops that systematically amplify emotionally charged and polarizing content because such material generates higher interaction rates than moderate or thorough information. Per Nechushtai et al. (2024), the recommendation systems on platforms like Facebook, Twitter and YouTube prioritize content similarity and user clustering, leading to the formation of ideologically homogeneous information environments. Research demonstrates that partisan content receives significantly more algorithmic promotion than neutral news sources, as divisive material triggers stronger emotional responses that translate into measurable engagement metrics (Kearney et al., 2025). The economic incentives

underlying digital advertising models reinforce these tendencies, as platforms generate revenue through sustained user attention rather than information accuracy or democratic discourse quality.

Inference from the empirical literature showed that aalgorithmic amplification operates through multiple mechanisms that systematically favor partisan content over balanced journalism in the American media ecosystem. Content recommendation algorithms use collaborative filtering techniques that assume users with similar past behaviors will prefer similar future content, effectively creating echo chambers where partisan viewpoints are continuously reinforced (Lee, 2024). The algorithmic prioritization of recent, trending and highly engaging content disadvantages traditional journalism practices that emphasize verification, context and measured analysis. Tsekhmeistruk (2024) noted that Partisan content producers have adapted their strategies to exploit algorithmic preferences by employing sensationalized headlines, emotionally provocative imagery and rapid publication schedules that align with platform optimization criteria. Studies of Facebook's news feed algorithm reveal that political content with strong partisan framing receives up to 600% more distribution than equivalent moderate content (Piccardi et al., 2024). The speed and scale of algorithmic distribution mean that partisan narratives can achieve widespread circulation before fact-checking mechanisms or editorial oversight can respond effectively.

Diaz Ruiz (2025) underscored that the structural design of digital platforms creates asymmetric information environments that disproportionately benefit partisan content creators over mainstream media outlets. Platform algorithms favor content that generates immediate engagement over content that provides long-term informational value, systematically disadvantaging in-depth reporting and investigative journalism that require sustained attention. The democratization of content creation through social media has enabled partisan actors to bypass traditional gatekeeping mechanisms, which allows ideologically motivated sources to compete directly with established news organizations for audience attention. Metrics-driven content curation has transformed the American information landscape by creating economic incentives for polarization, as moderate positions generate less algorithmic amplification and therefore less advertising revenue (Ahmed et al. 2025). The concentration of digital media consumption within a small number of major platforms means that algorithmic decisions made by technology companies have unprecedented influence over public discourse and democratic participation. Understanding these architectural features is essential for addressing information polarization, as technical design choices embedded within platform algorithms have become critical determinants of political communication patterns in contemporary American democracy.
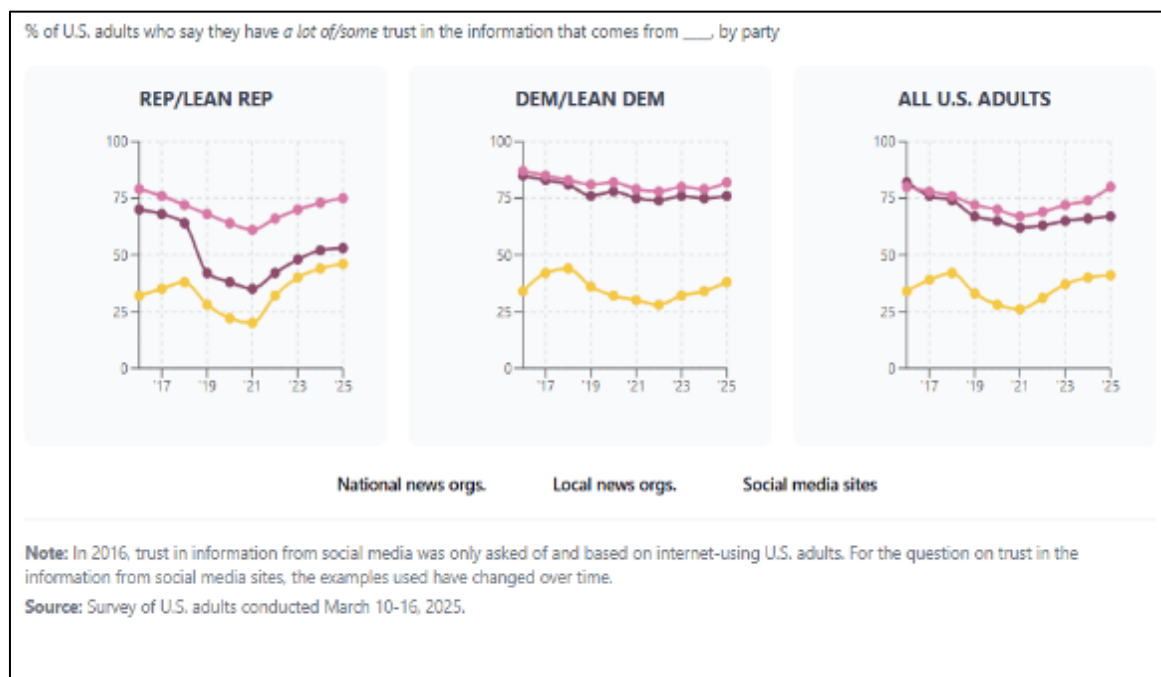

Figure 2 Republicans have become more likely since 2024 to trust information from news organizations, social media sites

The figure above illustrates a striking partisan divergence in media trust over the past decade, with Republicans showing significantly more volatility in their confidence in news sources compared to Democrats. Most notably, Republican trust in national news organizations plummeted from 70% in 2016 to a low of 35% in 2021, before recovering to 53% by 2025; however, Democratic trust remained relatively stable around 75-85% throughout the same period. The data reveals that local news organizations consistently maintain higher trust levels than national outlets across all political affiliations, with Republicans trusting local news at 75% and Democrats at 82% in 2025. Social media trust remains consistently low across the political spectrum, hovering around 40-46% for Republicans, 34-38% for Democrats and 41% overall by 2025, which suggests broad scepticism toward social platforms as reliable information sources. The Republican recovery in trust since 2024, as highlighted in the chart's title, represents a partial reversal of the dramatic decline that occurred during the Trump presidency and early Biden administration, though trust levels have not returned to their 2016 peaks.

2.5. USA Institutional Responses and Democratic Resilience: Fact-Checking, Media Literacy and Platform Governance

The institutional response to misinformation in the United States has centered on three primary mechanisms that collectively attempt to preserve democratic integrity through information verification and platform accountability (Tennova, 2020). Fact-checking organizations have emerged as crucial arbiters of information accuracy, though research demonstrates their effectiveness depends heavily on public trust in these institutions as impartial evaluators. Social media platforms have implemented content moderation systems that embody what scholars describe as a libertarian approach to truth verification, which assigns users rather than authoritative fact-checkers the primary role in determining information credibility (Wehby et al. 2020; Ukkola, 2025). The existing literature also noted that media literacy initiatives have gained prominence as educational interventions designed to enhance citizens' capacity to critically evaluate information sources, with studies indicating that individuals with higher analytical skills demonstrate greater resilience against false narratives.

The governance of digital platforms has evolved into a complex regulatory landscape where technology companies face increasing pressure to balance free expression with misinformation prevention. Recent platform policy changes, including Meta's 2025 decision to reduce fact-checking in favor of community-driven content evaluation, reflect the ongoing tension between automated content moderation and human oversight, with companies acknowledging significant error rates in their enforcement mechanisms (Vinhos and Bastos, 2025; Wischnewski, 2022). The OECD framework emphasizes that while false information may not be illegal, it poses substantial risks to democratic processes by contributing to polarization and undermining trust in institutions. Government responses have focused on promoting information integrity without direct censorship, which relies instead on transparency requirements, platform accountability measures and support for independent verification organizations.

The resilience of democratic institutions ultimately depends on citizens' access to reliable information necessary for meaningful political participation, including the ability to evaluate candidates, assess policy proposals, and engage in informed civic discourse (Monsalve et al. 2024). Democracy's epistemic foundation requires a shared knowledge base among citizens, particularly trust in electoral processes and access to accurate information for policy debates (Tenove, 2020).

However, experts recognize that misinformation's power derives not from factual content but from emotional responses it generates, which suggests that institutional responses must address both cognitive and affective dimensions of information processing. The effectiveness of these institutional mechanisms remains contested, with ongoing debates about whether current approaches adequately protect democratic discourse, however preserving fundamental rights to free expression and open debate.

3. Discussion

This research elucidates the complex pathways through which disinformation campaigns systematically undermine democratic governance in the United States, which operates through interconnected mechanisms that extend beyond simple information distortion to encompass broader epistemic disruption. The empirical findings demonstrate that disinformation functions as a strategic instrument of political warfare, which is deliberately designed to fragment the shared knowledge base upon which democratic deliberation depends. The convergence of algorithmic amplification systems, psychological predispositions toward confirmation bias and the structural affordances of digital platforms creates what scholars term an "infodemic environment" wherein false narratives achieve viral propagation. However, accurate information struggles for comparable reach. This phenomenon manifests most acutely in the formation of

ideologically segregated information ecosystems that reinforce existing partisan identities but actively prevent the cross-cutting exposure essential for democratic compromise and consensus-building.

The paper's analysis of institutional responses reveals a fundamental mismatch between the scale and sophistication of disinformation threats and the capacity of existing democratic safeguards to address them effectively. Traditional mechanisms of information verification, including journalistic fact-checking and educational media literacy initiatives, operate on temporal and distributional scales that cannot compete with the instantaneous global reach of algorithmically optimized false content (Wihbey et al, 2020). The recent policy shifts by major technology platforms toward community-based content moderation, ostensibly promoting user agency, paradoxically exacerbate the problem by distributing verification responsibilities to populations already subject to polarized information. The fragmentation of regulatory authority across federal agencies, state governments and private sector entities has precluded comprehensive policy coordination, which results in what this research characterizes as "governance gaps" that malicious actors systematically exploit.

The implications of these findings extend beyond immediate concerns about electoral integrity to encompass fundamental questions about the sustainability of democratic institutions under conditions of persistent epistemic instability. The erosion of shared factual foundations necessary for democratic participation threatens not only individual voting decisions but the broader legitimacy of democratic outcomes, as citizens increasingly question the validity of information sources, electoral processes and governmental authority itself. This research suggests that addressing disinformation-driven polarization requires a paradigmatic shift from reactive content moderation to proactive information architecture design, which emphasizes transparency in algorithmic systems, investment in civic education infrastructure and the development of institutional frameworks capable of preserving both information integrity and democratic pluralism. Without such comprehensive reform, the United States faces the prospect of what political scientists' term "competitive authoritarianism," wherein democratic forms persist while their substantive content is hollowed out by systematic manipulation of the information environment upon which democratic choice depends.

4. Conclusion

This study establishes a comprehensive framework for understanding the intricate mechanisms through which disinformation undermines democratic stability by catalysing and intensifying political polarization in contemporary American society. The research demonstrates that the erosion of democratic norms occurs not through singular pathways but through the complex interplay of psychological vulnerabilities, technological amplification systems and weakened institutional safeguards that create cascading effects across the political landscape. The findings underscore that combating information disorder requires moving beyond reactive approaches toward systemic interventions that address root causes of democratic fragility, including the restoration of institutional trust, enhancement of civic education and implementation of technological reforms that prioritize democratic values over engagement metrics. Although the digital age presents unprecedented challenges to shared truth-making and democratic discourse, this paper reveals that the preservation of American democracy depends fundamentally on our collective capacity to rebuild the social and institutional infrastructure necessary for constructive political dialogue, evidence-based decision-making and legitimate governance.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Ahmed, T. N., Mustafa, N. N., Ahmed, R. K., Saeed, M. S., Ali, A. Q., & Qadir, K. A. (2025). The Impact of Digital Technologies on Journalistic Integrity: An Analysis of Clickbait, Algorithmic Influence and Societal Consequences. Asian Journal of Education and Social Studies, 51(6), 566-580.

[2] Baldassarri, D., & Page, S. E. (2021). The emergence and perils of polarization. Proceedings of the national academy of sciences, 118(50), e2116863118.

[3] Baliñas Pérez, D. (2024). On Political Polarization: Conceptualizing Affective Polarization and Exploring its Causes from a Multi-method Approach.

[4]   Beaumont, E., & Beerbohm, E. (Eds.). (2024). Civic Education in Polarized Times: NOMOS LXVI (Vol. 38). NYU Press.

[5]   Benkler, Y., Faris, R., & Roberts, H. (2018). Network propaganda: Manipulation, disinformation, and radicalization in American politics. Oxford University Press.

[6]   Blascovich, G. M. (2017). Threat and Polarization: Social Identity and Bias among American Partisans (Doctoral dissertation, Stanford University).

[7]   Citrin, J., & Stoker, L. (2018). Political trust in a cynical age. Annual review of political science, 21(1), 49-70.

[8]   Cushman, F. (2020). Rationalization is rational. Behavioral and Brain Sciences, 43, e28.

[9]   Dahlgren, P. (2018). Media, knowledge and trust: The deepening epistemic crisis of democracy. Javnost-The Public, 25(1-2), 20-27.

[10]   Diaz Ruiz, C. (2025). Disinformation on digital media platforms: A market-shaping approach. new media & society, 27(4), 2188-2211.

[11]   Ferrara, E. (2020). Bots, elections, and social media: a brief overview. Disinformation, misinformation, and fake news in social media: Emerging research challenges and opportunities, 95-114.

[12]   Kearney, A., Poredi, N., Shelton, J. A., Akcinaroglu, S., Karakoc, E., Tran, T., & Chen, Y. (2025, May). Echoes amplified: a study of AI-generated content and digital echo chambers. In Disruptive Technologies in Information Sciences IX (Vol. 13480, pp. 164-202). SPIE.

[13]   Kingzette, J. (2021). Partisan Animosity and the Civic Culture: Examining the Causes and Consequences of Affective Polarization in the United States. The Ohio State University.

[14]   Kleinfeld, R. (2023). Polarization, democracy, and political violence in the United States: what the research says.

[15]   Kruglanski, A. W., Dechesne, M., Orehek, E., & Pierro, A. (2018). Three decades of lay epistemics: The why, how, and who of knowledge formation. In The Motivated Mind (pp. 19-59). Routledge.

[16]   Lee, M. (2024). Is Polarization an Inevitable Outcome of Similarity-Based Content Recommendations?--Mathematical Proofs and Computational Validation. arXiv preprint arXiv:2412.10524.

[17]   Lee, P. C. (2024). User engagement in digital curation: A systematic review and synthesis. Journal of Librarianship and Information Science, 09610006241291607.

[18]   Livingston, S., & Bahador, B. (2025). Propaganda feedback loops as communication rituals: Hate speech on talk radio. Media, Culture & Society, 01634437251331809.

[19]   Marino, E. B., Benitez-Baleato, J. M., and Ribeiro, A. S. (2024). The polarization loop: How emotions drive propagation of disinformation in online media—The case of conspiracy theories and extreme right movements in southern Europe. Social Sciences, 13(11), 603.

[20]   Marsden, C., Meyer, T., and Brown, I. (2020). Platform values and democratic elections: How can the law regulate digital disinformation?. Computer law and security review, 36, 105373.

[21]   Monsalve, L. F. D., Valladares, C. P. N., and Díaz, J. S. (2024). Relationship between political participation and community resilience in the disaster risk process: A systematic review. International Journal of Disaster Risk Reduction, 104751.

[22]   Nechushtai, E., Zamith, R., and Lewis, S. C. (2024). More of the same? Homogenization in news recommendations when users search on Google, YouTube, Facebook, and Twitter. Mass Communication and Society, 27(6), 1309-1335.

[23]   Papp, J. T. (2024). Towards a More Informed Public: The Role of Media (Pluralism) in Mitigating Filter Bubbles. Journal of Information Ethics, 33(2).

[24]   Phillips, L., Dowling, C., Shaffer, K., Hodas, N., and Volkova, S. (2017). Using social media to predict the future: a systematic literature review. arXiv preprint arXiv:1706.06134.

[25]   Piccardi, T., Saveski, M., Jia, C., Hancock, J. T., Tsai, J. L., and Bernstein, M. (2024). Social Media Algorithms Can Shape Affective Polarization via Exposure to Antidemocratic Attitudes and Partisan Animosity. arXiv preprint arXiv:2411.14652.

[26]   Pilditch, T. D. (2017). Communicated beliefs: The interplay of evidence and truth values in erroneous belief acquisition and maintenance (Doctoral dissertation, UCL (University College London)).

[27] Robison, J., and Mullinix, K. J. (2016). Elite polarization and public opinion: How polarization is communicated and its effects. Political Communication, 33(2), 261-282.

[28] Shelley, T. (2022). Modern Liberty, Partisanship, and Identity: On Mark Lilla and Yuval Levin. Canadian Review of American Studies, 52(2), 206-228.

[29] Tenove, C. (2020). Protecting democracy from disinformation: Normative threats and policy responses. The International Journal of Press/Politics, 25(3), 517-537.

[30] Tsekhmeistruk, R. (2024). QUANTIFYING ALGORITHMIC BIAS IN NEWS RECOMMENDATIONS: METHODOLOGIES AND CASE STUDIES. Scientific Journal of Polonia University, 66(5), 251-259.

[31] Ukkola, A. (2025). The influence of social media recommendation algorithms on opinion polarization.

[32] Vinhas, O., and Bastos, M. (2025). The WEIRD governance of fact-checking and the politics of content moderation. New Media and Society, 27(5), 2768-2787.

[33] Vosoughi, S., Roy, D., and Aral, S. (2018). The spread of true and false news online. science, 359(6380), 1146-1151.

[34] Webster, S. W., and Abramowitz, A. I. (2017). The ideological foundations of affective polarization in the US electorate. American politics research, 45(4), 621-647.

[35] Whyte, C. (2020). Cyber conflict or democracy "hacked"? How cyber operations enhance information warfare. Journal of Cybersecurity, 6(1), tyaa013.

[36] Wihbey, J. P., Kopec, M., and Sandler, R. (2020). Informational Quality Labeling on Social Media: In Defense of a Social Epistemological Strategy. Yale JL and Tech., 23, 153.

[37] Wischnewski, M. (2022). Misinformation on social media: Investigating motivated reasoning through an identity-protection model (Doctoral dissertation, Dissertation, Duisburg, Essen, Universität Duisburg-Essen, 2021).