The Public Influence Triad
Understanding and Addressing Polarization in our System of Public Discourse

Eric Saund
eric.saund@gmail.com
v10
2021/04/13

Introduction: The Polarization Problem

By now there is broad agreement that public discourse is severely broken in America. Common characterizations are: political polarization; tribalization; political sectarianism; epistemic polarization; echo chambers; filter bubbles; alternative realities. In a democracy, this is a problem. The good thing about a democracy is that governance is designed to follow the will of the people. The weakness of a Democracy is that if the wills of most people are not reasonably coherent and consensual, governance breaks. Today, Americans are divided into tribes of mutual disdain and suspicion that has destroyed trust in the system itself. We suffer widespread dysfunction, pockets of rebellion, and episodes of chaos.

In order to consider remedial policy proposals, we must accurately diagnose the problem. We need a shared mental model for the factors and forces of influence that govern public common knowledge, belief, and opinion. Our goal should be a system that encourages diversity of thought and belief, yet modulates discourse so as to maintain cohesion and common ground.

Three central categories of influence- and belief-actor define how public opinion is formed and held: *Political Leaders, Platformed Thought Leaders,* and the *Grassroots Thought Ecosystem*. Interactions among these foundational elements sustain a well-functioning democracy, but also a broken one. Then, two other categories of influencer come into play, *Moneyed Interests,* and *Malevolent Agitators*. By charting the main influence pathways among these actors, we build a common framework that situates current observations, criticism, and policy recommendations addressing the polarization problem. Dysfunctions are systemic. Each of the elements operates according to local incentives and constraints, but the collective bad outcome that we observe is self-reinforcing. Importantly, it is unlikely that any single policy intervention can significantly change the system dynamics on its own; multiple policy levers will need to be applied in concert with one another.
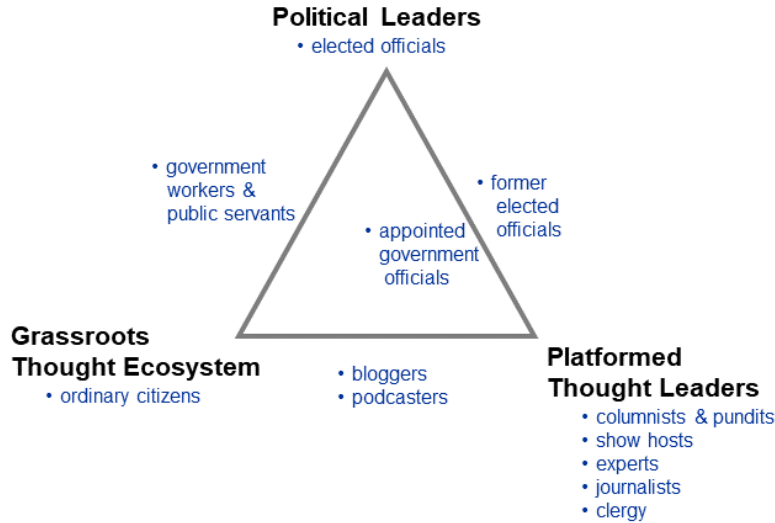
I. Framework

The Public Influence Triad of Actors

Political Leaders
In a democracy, political leaders are primarily elected officials. The United States has about half a million elected officials across local, state, and national levels. These are the people who are make major governmental decisions, and are held accountable for them.

The Public Influence Triad
(United States)

Political Leaders
• elected officials

• government workers & public servants

• former elected officials

• appointed government officials

Grassroots Thought Ecosystem
• ordinary citizens

• bloggers
• podcasters

Platformed Thought Leaders
• columnists & pundits
• show hosts
• experts
• journalists
• clergy

Platformed Thought Leaders

Platformed Thought Leaders have ways to convey their ideas to a relatively large number of others through traditional media such as broadcast news and opinion, print, digital publications, and speaking venues. This category includes television and radio hosts, opinion columnists, pundits, and commentators. Academics, relevant experts, business leaders, and other prominent individuals are given platform voice according to their expertise and audience interest. The clergy propagates and reflects values-based judgments about current affairs to their flocks. Because news, opinion, and editorial selection and filtering weigh into the content carried on platforms, journalists and editors fall under this category as well.

Platforms range in size and reach, from broadcast networks and major current affairs magazines down to local newspapers and small newsletters. The number of platformed thought leaders in the United States is probably in the tens to hundreds of thousands, depending on how you count.

Grassroots Thought Ecosystem

Most opinion about public affairs is carried in the minds of ordinary citizens---the grassroots. Folks without a platform discuss topics with each other in informal person-to-person settings, and recently, in digital social network spaces. Based on election turnout, we can estimate that about 150 million Americans participate in this category to at least a minimal degree of awareness and concern about current affairs and public policy.

Falling between these corners, the sides of the triad represent people who bridge the major categories.

Bottom side:

Ordinary citizens can sometimes address larger audiences on formal platforms. Anyone can call in to a talk show, write a letter to the editor, or speak up at a town meeting. More significantly, the digital age has democratized self-publishing on blogs, podcasts, and social network groups. Thousands or perhaps millions of outspoken, ambitious, and talented people migrate along the zone between small-scale grassroots communication and large-reach established platforms.

2

Left side:

Nearly 24 million people work for government in some capacity. In terms of public discourse, these people are not responsible for shaping or responding to public opinion, but they are more directly connected to public policy and decisions than the ordinary citizen.

Right side:

A significant number of former government officials fall somewhere between the Political Leader corner and Platformed Thought Leaders. Their public profiles and expertise earn them both influence with serving elected officials, and access to established media platforms.

Inside:

High profile appointed government officials such as cabinet secretaries, agency leads, and commission members bridge all three principal categories of public thought participation and influence. Though not elected, they engage in and are responsible for government policy, while their status gains them voice on media platforms.

Belief Influence Pathways

Consider next the pathways of dialog and influence among the actor categories, drawn in red.

The Public Influence Triad
(United States)
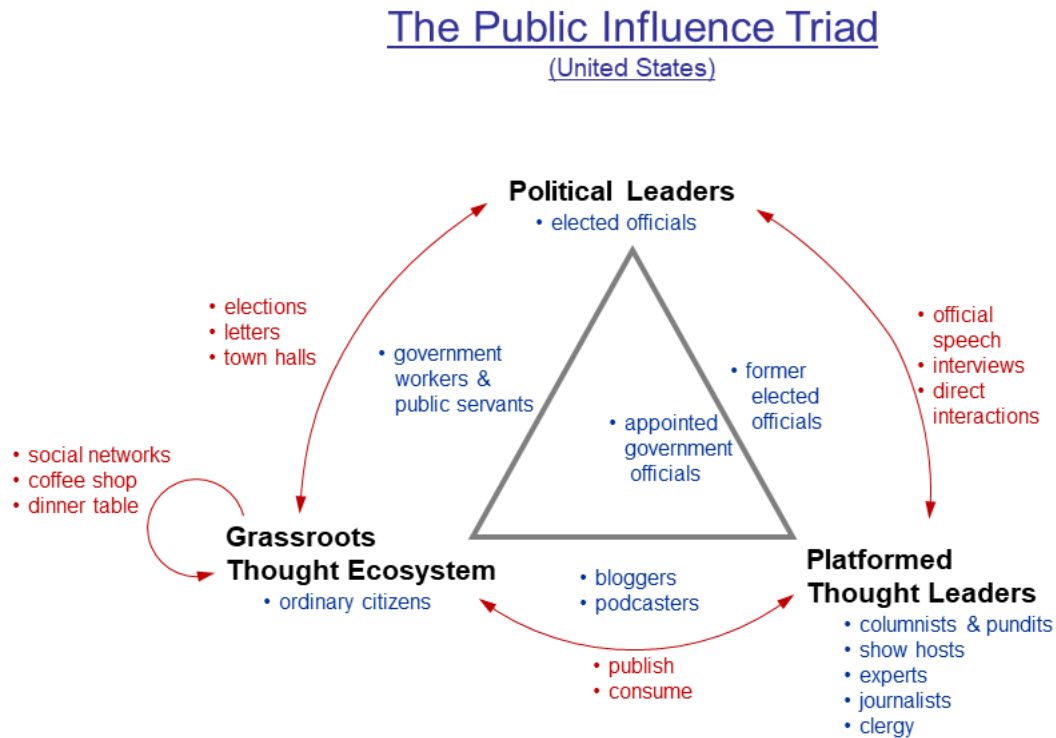
Ideally, Political Leaders engage in a two way transaction with the public. They give speeches and hold town hall meetings. The public communicates back through meetings, letters, and petitions. In a democracy, the Grassroots Thought Ecosystem selects leaders and policies by voting on individual offices and policy measures.

Dialog between the Grassroots and Platformed Thought Leaders takes place through the routine transactions of publishing and consuming editorial content. Platforms publish and promote their wares, which include entertainment, news, analysis, and opinion. Oftentimes these are mixed together. These products are consumed by readers and viewers, who in turn decide which platform offerings to purchase, tune into, follow, and view online.

Meanwhile, the Grassroots Thought Ecosystem is churning away on its own on digital social networks and in-person networks---the coffee shop, the dinner table, the PTA meeting, the fishing trip.

Finally, consider the influence channel between Political Leaders and Platformed Thought Leaders. Obviously, Political Leaders are highly attuned to the open media---especially what is said about themselves. The official and unofficial words and deeds of Political Leaders generally command media attention, gaining them platform exposure by default. In addition, these actors give interviews to journalists. In the reverse direction, outside of their authored publications, Platformed Thought Leaders carry direct sway with Political Leaders sometimes through private conversations, but also commonly via interactions at social gatherings, conferences, panel discussions, current affairs forums, and other limited-audience venues.

## Influence Flows in an Ideal Democracy

The three corners of the Public Influence Triad and the pathways among them tell a story of how the public discourse system is supposed to work: The People learn about the views and deeds of candidate and office-holding officials through reporting and analysis by the platformed media--the fourth estate, as it were. Through selective consumption by the Grassroots, the "marketplace of ideas" guides Platformed Thought Leaders toward well-supported and well-considered reporting and opinions. The Grassroots segment reflects and formulates its beliefs through myriad local, bottom-up belief formation processes. Ordinary citizens learn about candidate Political Leaders directly and through the media platforms, then select them through elections. While holding office, Political Leaders continue to exchange views with the public. In order to be reelected, they must be responsive to the wishes and beliefs of the public. Meanwhile, Platformed Thought Leaders and Political Leaders communicate directly, sometimes in the open, sometimes in more secluded or obscure spaces, and sometimes in backchannels. Through platform access, Political Leaders and private sector thought leaders can nudge the public toward understandings and policy views that the Grassroots Thought Ecosystem would not arrive at on its own.

Under this framework, various modes of successful versus faulty dynamics can take hold. For example, polarization and alternative social-political realities emerge when the communication and influence pathways diverge into competing and mutually incompatible belief systems. Commentators have identified a number of indicators and factors that contribute to this breakdown. These include:

- A supply and demand dynamic whereby Platformed Thought Leaders amplify narratives that exploit cognitive biases in the Grassroots audience, driving bifurcated affinity groups to seek more extreme content validating their respective sides' worthiness. Platforms compete for audience share by crafting belief structures that further accentuate emotional appeal and demonize opponents. The Thought Leaders they carry become gladiators in an epic battle for hearts, minds, and souls.

- Within the Grassroots Thought Ecosystem, social media elevates extreme content because this increases engagement metrics.

- The democratization of platform access through blog and podcast channels increases the reach of self-publishers, while eliminating the reins of editorial judgment that traditionally constrained larger established platforms. (Bottom side of the triangle.)

- Political Leaders are elected who appeal to one Grassroots base or the other but do not span differences. In similar supply-and-demand manner as the Grassroots/Platformed Thought Leader relationship, Political Leaders justify extreme views and encourage polarization among the electorate in accordance with one side or the other of divergent and incompatible Grassroots Thought Ecosystems' demands. Divisive rhetoric amplifies engagement and commitment on the part of the respective poles of the electorate.

- Political Leaders oblige themselves to political parties that elevate the interests of partisan dominance over policy. Cooperation with the other side, or even acknowledgement of their merits, is treated as betrayal.

- Political Leaders and Platformed Thought Leaders selectively promote and associate with one another according to political tribe. This creates insular echo chambers of reinforcing belief and opinion among the axis of the two Leadership segments.

The Public Influence Triad
(United States)
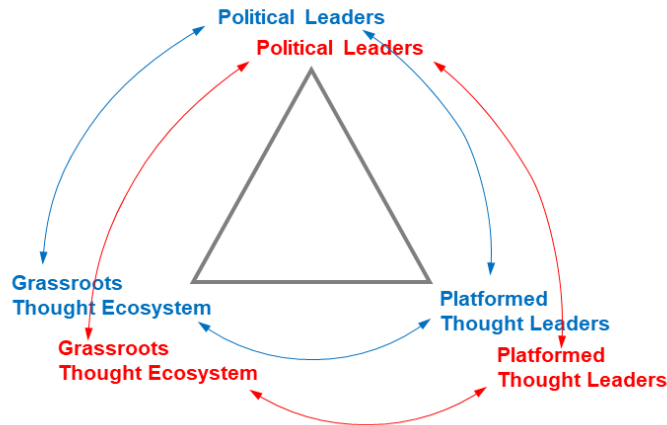

Alternate political realities.

By contrast, recent history has featured modes of system dynamics that favored pluralistic competition but avoided the degree of epistemic breakdown we are witnessing today. It's not that extreme views were once excluded across the three segments of belief and influence holder, but rather that extremism did not dominate. Cross-aisle rivalry has not always been toxic and it does not need to be. Much has been written about the conditions that encourage modulated and civil discourse and discourage polarization. These include:

- Platformed Thought Leadership:
    - common reporting of basic facts across news and information sources
    - standards of demeanor among commentators

- Grassroots Thought Ecosystem:
    - norms of behavior in public discourse
    - education about basic world knowledge, civics, the devices of rhetoric, skeptical awareness of tricks, tomfoolery, and manipulation
    - absence of means and incentives to share inflammatory rhetoric, especially in digital social media
    - trust in institutions

- Political Leaders
    - norms of behavior and conduct by political officeholders
    - incentives for bipartisan cooperation and credit-sharing

Below, we consider policy options that might inject stabilizing factors into various parts of the system. But first, let us consider two additional categories of thought and influence actor who play increasingly significant roles in the digital era.

The Public Influence Triad
(United States)
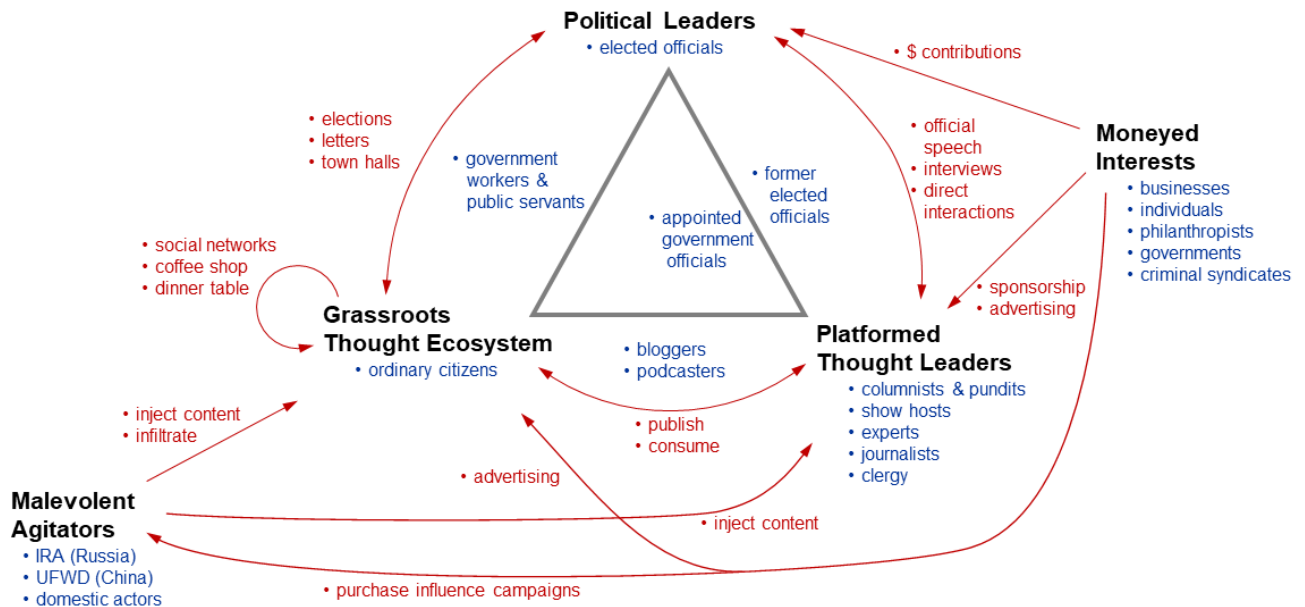
Moneyed Interests

Moneyed interests are actors who possess the means and motivation to deliberately influence the public discourse. The political system generates policy that these interests care about, so of course they want to shape the terms of discussion. These interests include businesses, wealthy self-interested individuals, philanthropists, criminal syndicates, and government entities.

The ways by which Moneyed Interests purchase influence are well-established. For one thing, they give money to Political Leaders, primarily to support campaigns, but sometimes for private enrichment. Moneyed Interests sponsor advertising on formal platforms, oftentimes in alignment with the Platformed Thought Leaders that they host. More recently, Moneyed Interests purchase advertising on social media sites that micro-target to different belief and demographic segments in the Grassroots Thought Ecosystem.

Moneyed interests have always been major outside participants in the idealized Public Influence Triad. Without counterbalancing instruments, their means and methods tend to self-perpetuate and grow over time. Accordingly, our system has introduced controls such as contribution limits, truth in advertising laws, and conflict of interest rules. Still, most agree that Moneyed Interests play an outsized role in shaping public discourse.

Malevolent Agitators

A fifth category of influence actor has attracted notice and alarm. Malevolent Agitators conduct influence campaigns designed to exploit weaknesses in our discourse dynamics. Propaganda and influence have been potent devices for political and military action throughout history. Techniques and methods have by now been honed and weaponized for the information age. These include purchasing of advertising and editorial content among Platformed Thought Leaders, and sponsoring fake or legitimate voices in the grassroots social media arena and in the democratized blogosphere and podcasting domains. Organized Malevolent Agitators sometimes pose as everyday citizen members of the grassroots. Injection of deliberate disinformation, or even selected truths framed by distorted narratives, leads to confusion and propagation of misinformation on the part of even well-meaning participants among the Grassroots Thought Ecosystem.

While participants throughout the system may display malign motives and methods, organized Malevolent Agitator brigades require resources to operate. These are supplied by Moneyed Interests, which can include governments, businesses, and criminal syndicates for hire. Among the most pernicious and successful Malevolent Agitators are the Russian Internet Research Agency, and the Chinese United Front Work Department. When deployed in conjunction with cyberspace operations, Malevolent Agitators' toolkits can include hack-and-release campaigns and blackmail.

A growing academic and public investigation discipline has arisen in recent years aimed at identifying and tracking disinformation and information operations. Their research has identified many instances and templates by which Malevolent Agitators have destabilized traditionally sound patterns for belief formation within the Public Influence Triad. By injecting bad faith content into communities and communication channels, Malevolent Agitators have exacerbated the forces that break trust and common values necessary for healthy and open discussion of issues that matter.

Dynamic Attractor

This framework suggests that there is no single cause for our current plight of alternative realities in public discourse, and therefore no single solution. We got to this point not by design, but through the rational behaviors and adaptations of individual localized actors and communities.

"Rational behaviors and adaptations," does not mean that people are rational, logical thinkers. Human cognition is driven by a mixture of logical thought, shortcut heuristics, and emotional reaction. Beliefs and opinions about complex topics are seldom supported by all available facts. Everyone's knowledge and experience is limited.

"Rational behaviors and adaptations" instead means that people maintain goals and aspirations, they hold values, they care about a variety of concerns to different degrees, and they respond to incentives. Political and Thought Leaders generally enjoy the power and prestige of their positions, and they respond to market forces from their constituencies to maintain them. Ordinary citizens vary widely in knowledge and interest in public affairs, in their idealism versus cynicism, and in their dedication to self, family, community, religion, philosophy, or nation. Incentives for accepting, adapting, and promoting beliefs about topics of the day arise from approval of one's peers, from the satisfaction of constructing and sharing explanations with one's community, and simply from the attention gained by spouting off. All categories of participant in the Public Influence Triad play both proactive and reactive roles.

Thus, the Public Influence Triad describes a system of interacting parts. The emergent outcome---coherent function or else breakdown---arises from setpoints and parameters of the system elements. For example, different content moderation policies can lead to the promotion or demotion of different tenors of discussion within the Grassroots Thought Ecosystem. Such an adjustment may or may not propagate to a desirable downstream effect, such as, that the aggregate market behavior of ordinary citizens steers Platformed Thought Leaders away from extreme content and toward thoughtful moderation.

This conceptual framework supports thought experiments and imagination for policy measures and adjustments that could tweak or transform the emergent dynamics. It would be far preferable to go beyond a verbal and diagrammatic description, to build computational simulation models in which we could perform controlled experiments. What exactly would be the effect, for example, of re-introducing the Fairness Doctrine that once governed the broadcast segment of Platformed Thought Leaders? That level of technically powered system modeling lies in the future.

We would, however, advocate one hypothesis in particular. Namely, the system is, in the terms of complex systems theory, a *dynamic attractor*. That means that the belief and influence dynamics that have brought on today's political polarization are self-reinforcing. Tweaking one or a few parts of the system will not cure the polarization problem. Instead, we will need to make policy adjustments across many of the elements and interaction pathways. Can we prove this hypothesis? No. But it does seem as if we are stuck in a rut. Everybody needs to get out of the car and push in the same direction.


II. Policy Considerations

Policy Objectives

We often speak of creating a level playing field where all ideas have equal opportunity to compete and take hold in people's minds. The marketplace of ideas should be open and accessible to all. Yet, we know that other kinds of markets, notably financial markets, require rules and constraints in order to fulfill their ideals. If left unfettered, the powerful devise means to monopolize information, set their own rules, and ensure that they come out ahead. Therefore, we acknowledge the need for some degree of market regulation.

Similarly, a level playing field for the marketplace of ideas is nice in theory, but in practice, that's not what we have. The playing field is warped to favor some types of "ideas" over others. This is in part because of human cognitive biases, limited attention, and the structural organization of the Public Influence Triad.

Skeptics of policy proposals justifiably express concern about tilting the playing field toward one political side or the other. We should all agree that out of fairness and humility, this is not a desirable goal. The goal is to correct the *curvature* of the playing field.

The Public Discourse Playing Field

Tilted Left

Tilted Right

Dome (divisive)

Bowl (unifying)
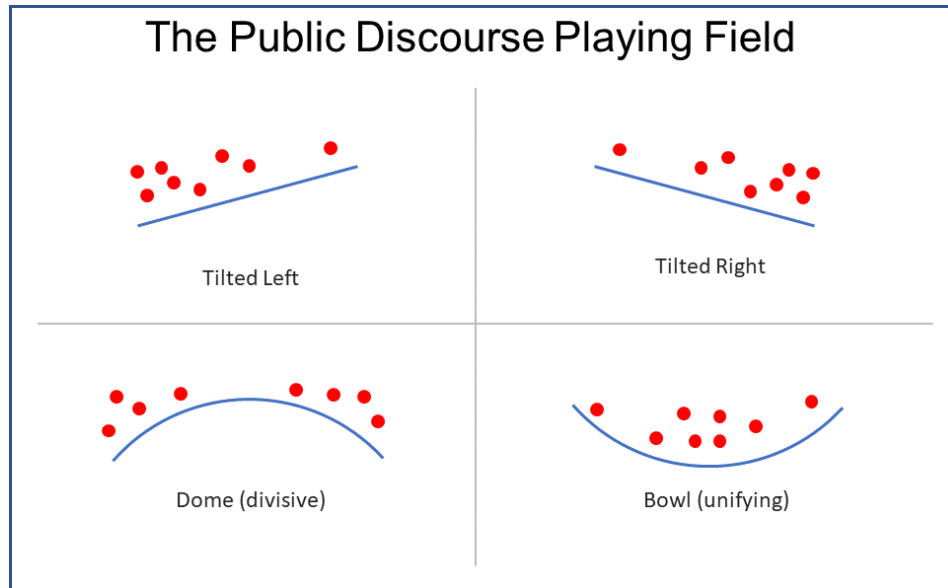
Currently, the public influence system is warped to take a concave downward, dome shape. The belief formation system disfavors moderation, even-handedness, modesty in rendering judgment, honesty about uncertainties, and listening to both sides. This warpage drives people away from the center and toward the extremes.

Instead, policy levers can bend the playing field into a bowl shape. The bias becomes then against extreme viewpoints, and toward concurrence. The "middle" isn't fixed; that depends on particular issues and can shift leftward or rightward as consensus takes it. A convergent dynamic doesn't mean that everyone is supposed to agree. That would defeat the purpose of pluralism. Instead, a gently sloping bowl preserves diversity of thought and exploration of alternative ideas, but it requires reaching to the common ground in order to recruit others. Outlier positions are by no means prohibited, they are simply not favored to capture everyone's minds by the default system mechanics.


Policy Catalog

Below is an admittedly incomplete catalog of existing policy proposals designed to adjust incentives and constrain behaviors away from political tribalization, and steer public discourse toward coherence and common cause. The Public Influence Triad framework provides conceptual language in which to situate and evaluate not only these, but also other policy remedies overlooked in this article, or yet to be conceived, that would help shift it to a better dynamical mode.

These policy proposals emphasize the philosophy of openness, freedom of thought, and freedom of expression that have historically sustained the strength and vibrancy of liberal democracies.   Closed and authoritarian societies may possess similar centers of influence and belief evolution across venues of popular thought, but they adopt very different strategies that emphasize central controls placed *by* Political Leaders *onto* Platformed Thought Leaders and the Grassroots Thought Ecosystem. We must develop viable alternatives to authoritarian methods.

Contributions of comments, suggestions, and additional policy ideas are welcome.

1 Defend Against Malevolent Agitators
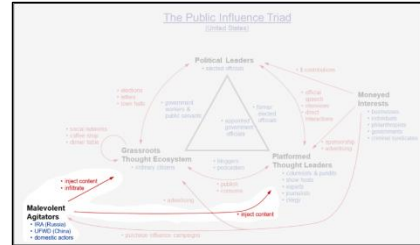
1.1 *Detect and expose influence operations.* Influence operations are designed to disrupt and redirect public discourse through deliberate manipulative methods. They may be foreign or domestic in origin and coordination.

> 1.1.1 *Support investigators.* Provide financial support to academic, nonprofit, and government centers and agencies whose mission is to detect, track, and expose influence operations and the Malevolent Agitators behind them.
>   Examples:
>   - State Department Center for Global Engagement
>   - Carnegie Partnership for Countering Influence Operations (PCIO),
>   - Atlantic Council *DFRLab (Digital Forensic Research Lab)*
>   - Stanford Internet Observatory
>   - Harvard Shorenstein Center on Media, Politics, and Public Policy
>   - University of Washington Center for an Informed Public
>   - DARPA INCAS program: INfluence Campaign Awareness and Sensemaking

> 1.1.2 *Better data access for tracking misinformation.* Current proprietary methods, user privacy concerns, and concerns about liability impede social media platforms from sharing data with legitimate organizations charged with detecting and exposing disinformation, influence campaigns and the operators behind them. This situation can be improved through legislation, standards, and norms.
>   e.g. John Bowers, Jonathan Zittrain,
>   "Answering impossible questions: Content governance in an age of disinformation"
>   Harvard Kennedy School Misinformation Review, January 14, 2020

1.2 *Authentication of accounts and users.* Currently, fake users, phony news sites, and sock puppets, and bots proliferate as handy tools of Malevolent Agitators. Perversely, metrics of traffic and growth create counterproductive incentives against social media platforms verifying the authenticity of users and accounts. On the other hand, for some purposes such as whistleblowing and dissent, user anonymity is desirable.

> 1.2.1 *Unique Users.* Social media sites could be required or incentivized to authenticate users and accounts, and uniquely tie each online account with a single human person or appropriately registered automated service.

1.2.2 *Digital Identity*. Proposed Digital ID laws such as the "Improving Digital Identity Act" can enable user authentication while preserving data privacy and anonymity.

1.3 *Regulation of foreign activities.* Among the most threatening Malevolent Agitators are foreign governments. Counter-measures include prohibitions and transparency requirements about foreign content in social networks, in platformed media, and in sponsorship of bloggers and podcasters. (Ref. Center for Humane Technology):

1.3.1 *Foreign agents*. Mandate registration by foreign agents + penalties for state-sponsored foreign interference; Foreign Agents Registration Act (FARA - US Department of Justice).

1.3.2 *Foreign sponsored ads.* SHIELD Act: "Stopping Harmful Interference in Elections for a Lasting Democracy Act"; FADE Act: "Foreign Agent Disclaimer Enhancement Act".

1.4 *Ad and content targeting.* The ability to target advertisements and tailored content to individuals or groups on the basis of demographic, psychological, or belief profiles is widely recognized as a point of weakness enabling malevolent actors to create division in our society. Restrictions on such targeting apply not only to the Malevolent Agitators, but also commercial marketing and political persuasion professionals whose aims may be not malevolent, but merely self-interested in ways that exploit data against the greater good. Therefore, policies in this category apply to the Grassroots Thought Ecosystem and the Platformed Thought Leader segments as well. (Ref. Center for Humane Technology):

1.4.1 *User Privacy.* Enact and enhance privacy laws about gathering, exploiting, and brokering user data.

1.4.2 *Transparency in ad targeting.* Enact requirements for transparency in targeting.

1.4.3 *Regulation of ad targeting.* Eliminate or regulate behavioral advertising and microtargeting.

1.5 *Civil information defense.* Disinformation campaigns on social media involve large quantities of postings, articles, and links that bombard ordinary citizens and overwhelm their natural cognitive defenses. Policies in this category apply to the Grassroots Thought Ecosystem and the Platformed Thought Leader segments as well.

1.5.1 *Civics education.* Sponsor and promote civics education in:
   o history
   o how government really works
   o cognitive psychology
   o techniques of misinformation
   o rhetoric and persuasion

1.5.2 *Gamified Awareness.* Sponsor info-ops inoculation games, e.g. HarmonySquare, promoted by the Center for Global Engagement.

1.5.3 *Public Service Announcements.* Run PSA campaigns to educate the general public in disinformation, misinformation, techniques of propaganda, cognition-hacking.

    1.5.3 *"Elves" corps*. In Lithuania, a group of civilian volunteers, called "Elves", fight Russian disinformation online. This proposal would require careful study and design to avoid potential pitfalls with regard to civil liberties, prohibitions against government propaganda, financing, and governance.

1.6 *Deepfake defenses.* Deepfakes are synthesized images, audio, and video that pass for authentic depictions of people and their speech and actions. Deepfakes can potentially falsely impersonate leaders and deceive viewers about leaders' statements and actions.

    1.6.1 Require mandatory disclosure of deepfake content, e.g. the "Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act".

    1.6.2 Prohibit inauthentic content in political advertising, e.g. the "Deepfakes in Federal Elections Prohibition Act"; SHIELD Act: "Stopping Harmful Interference in Elections for a Lasting Democracy Act"

1.7 *Transparency in sponsorship and advertising.* See 2.1

2. Defang Moneyed Interests

2.1 *Sponsorship and advertising.* Moneyed Interests influence public discourse through advertising and sponsorship on established platforms and social media platforms. Depending on their motivations and concerns, this could bend the playing field either in either a convex (divisive) or concave (unifying) direction. Steps can be taken to guard against the likelihood of divisive effects.

    2.1.1 *Regulation.* Regulate advertising on platform media, e.g. "Honest Ads Act".

    2.1.2 *Transparency*. Require transparency in advertising and sponsorship for established media and on social media.

    2.1.3 *Ad and content targeting*. See 1.4. This applies to benign intent as well as Malevolent Agitators.

2.2 *Cost of maintaining elected office.* The high cost of modern political campaigns leads politicians to spend inordinate time courting Moneyed Interests. Most of the cost goes to advertising, which is very profitable for media platform owners. Spending of money by corporate Political Action Committees has been deemed freedom of speech by the Supreme Court. A variety of campaign finance laws have been enacted, ruled upon by the courts, and remain under consideration.

2.5 *Transactions with Malevolent Agitators.* Malevolent Agitators for hire are able to conduct persuasion and info ops using tools, techniques, and procedures on behalf of Moneyed Interests beyond acceptable marketing, advertising, and promotion campaigns that can be used in the open. Laws and regulations can be established to prohibit and curtail such transactions.

> 2.1.5 *Dark money.* Influence campaigns are often purchased covertly using money laundering and dark money techniques. Campaign finance reform and financial regulatory reform laws can be enacted as barriers.

2.6 *Antitrust.* When Moneyed Interests own or control media platforms that they earn money from, then a positive feedback loop is created that incentivizes them to create and promote content that persuades the Grassroots Thought Ecosystem to elect Political Leaders and policies that cement these Interests' dominance. This applies both to established broadcast-scale media companies that host Platformed Thought Leaders, and to social media companies that host user generated content and smaller scale influencers.

> 2.6.1 *Deconsolidate media empires*. Antitrust law has been applied to various degrees to curtail media monopolies. See also 4.11, *Local news media*.

> 2.6.2 *Deconsolidate social media platforms.* The dominant social media giants benefit from "network effects". Namely, the benefits they convey to user-customers are magnified by the number of users on the platforms. Furthermore, people lock-in to dominant service ecosystems due to habit, and because services mutually reinforce one another. The use of antitrust laws to counteract such self-perpetuating dominance is under current consideration.

>> 2.6.2.1 *Data portability.* Social media platforms exercise great control over users' social media lives, including control over friend postings, user demographic and interest profiles, and viewing histories. Laws and regulations could be established requiring portability of this data across service providers. This would open competition and could reduce dominance by incumbents. Note however that fracturing of social media platforms could backfire by enabling new entrants to pursue divergent belief system markets among the Grassroots Thought Ecosystem customer base.

3. Adjust the Environment for Platformed Thought Leaders

3.1 *Editorial Requirements on Content.* Free speech principles prohibit government regulation of content published on print, broadcast, and digital media platforms except under narrow terms and circumstances. Permissibility of measures designed to nudge or coerce content toward the perceived public good have

been debated at length; the debate could be revisited in light of modern appreciation that foibles of human cognition undercut the idealized "marketplace of ideas".

3.1.1 *Protection of children*. Children especially are susceptible to psychological manipulation and enticement into unhealthy cognitive traps. The Children's Television Act of 1990 is among the legislation and rulings that establish content regulation requirements based on audience age.

3.1.2 *Liability for content*. Libel laws and court rulings have established limits on the permissible dissemination of falsehoods. These may be subject to updating.

3.1.3 *Fairness Doctrine.* The Fairness Doctrine placed requirements on political balance of content broadcast on public airwaves. The polarization issue has revived discussion of whether some form of ideological mixing might be reestablished in an updated form to cover cable television, internet streaming, and other newer platform technologies employed by Platformed Thought Leaders.

3.2 *Sponsorship, advertising, and incentives.* See 2.1.

3.3 *Reputational scoring.* While content regulation is problematical, public discourse is better served when greater information is available about the reliability, trustworthiness, biases, and accountability of media platforms and the Platformed Thought Leaders they carry. Methods are available and can be further developed to establish and promote standards, then score sources and voices.

3.3.1 *Fact checking.* Fact checking services alone have been found to be generally ineffective at counteracting strongly held narratives and deeply rooted false beliefs. But to the extent that they do become trusted arbiters, they can shape discourse toward veracity and form the foundation for aggregated reputational scoring.

3.3.2. *Media Bias.* A number of media quality and bias assessment methods have been introduced. For example the "Media Bias Chart" scores news and information publishers by political skew and originality of news reporting. Such services can inform the public in ways that could diminish the influence of extremist, outrageous, inflammatory, and nonsensical sources.

3.4 *Microtargeting and narrowcasting.* As discussed 1.4, tailoring of advertising and content is a recent development of the information age enabled by large scale data collection about viewers and users, algorithms such as recommender systems, computing horsepower, and communications bandwidth. The policy options for addressing microtargeting by Malevolent Agitators apply as well to contemporary incarnations of broadcasting and internet delivery of targeted content on the part of wide access platforms.

3.5 *Publicly funded platforms.* Relieved of commercial pressures, publicly funded news and information channels have proven successful at providing content that encourages convergent discourse based on trusted common ground. Examples are PBS, CPB, NPR, and the BBC. These organizations have however come under political scrutiny and skepticism on the part of some who find their content to be politically biased toward liberal viewpoints. This option may therefore be politically fraught.

3.6 *Political activity by charitable groups*. Financial favoritism in the form of tax relief is accorded to religious institutions that qualify as 501(c)(3) charitable organizations. Churches promote values that often have political implications for both issues and candidates for elected office. The boundary between values advocacy and political advocacy, which would remove tax exemptions, is blurry. Because highly charged rhetoric has become commonplace in these venues, this subject bears revisiting for the sake of shaping political discourse toward moderation.

4. Adjust the Environment for the Grassroots Thought Ecosystem
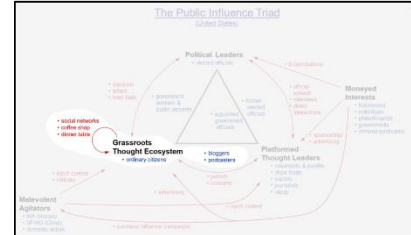
4.1 *Social media content moderation*. One of the most frequently cited factors in sectarian polarization has been the proliferation of untethered User Generated Content on social media. Two main mitigating approaches have been proposed, *content moderation*, which involves selectively blocking or removing certain content items and user accounts, and *content ranking and promotion*. Content moderation is a more drastic step because it can be viewed as censorship or violation of free speech rights.

> 4.1.1 *Liability protections.* Section 230 of the 1996 Communications Decency Act affords liability protections to social networks by not holding them responsible for user-generated content that they host, even when they selectively remove other user-generated content under good-faith exercise of content moderation policies. There is currently much debate about how Section 230 affects content moderation policies and thereby affects the shape of the public discourse playing field.

> 4.1.2 *Standards.* Part of the Section 230 and surrounding discussion of content moderation centers around the lack of existing standards for content evaluation and moderation. Proposals include:
> 4.1.2.1 Explicitly legislated standards.
> 4.1.2.2 Charging existing regulatory or standards bodies with developing standards for content moderation.
> 4.1.2.3 Setting up a new Digital Platform Agency with specific expertise and authorities for setting standards around internet content. See 4.9.

> 4.1.3 *Policy transparency*. Establish a requirement that social media platforms create and publish systematic standards for how they perform content moderation. This condition does not require the platform to expose technical details of algorithmic processes.

> 4.1.4 *Policy auditing*. Establish a requirement that independent agents perform audits of social media platforms' compliance with their own published policies. This will ensure that the content moderation results observed by end users and in aggregate by research organizations does in fact reflect policy objectives. This information is invaluable in consideration of adjustments to content moderation policies.

4.1.5 *Policy process transparency*. Policy process transparency goes beyond requiring that a content moderation policy be published, to requiring disclosure of how it is implemented, for example in an automated algorithm. This might be objected to by social media platforms for fear of exposing proprietary information.

4.2 *Content ranking and promotion.* The behavior of social network sites in ranking and promoting user generated content is also a strong factor in shaping the public discourse playing field. Investigators report people getting drawn further and further into extreme content because that's what algorithms determine drives user engagement. Similar policy options as discussed for content moderation in 4.1 apply here, with the exception of Section 230 liability protections which have not been applied to mere selective ranking of user-generated content (as opposed to total removal).

4.3 *Content flagging.* Content flagging, such as overlaying a notice that certain user-generated content is disputed, is another soft moderation device employed by social media platforms and implemented with algorithms. Similar policy options with regard to standards, policy transparency, policy auditing, and policy process transparency apply as in 4.1.

4.4 *User authentication and ID verification.* The policy options of 1.2 apply on social media not only to organized Malevolent Agitators, but also to small groups and lone wolves.

4.5 *Speaker sanctions and removal.* Social media platforms regularly ban user accounts that engage in repeated disruptive behaviors and posting of extremist content. However, financial incentives are in favor of policing mildly, and policies for blocking users are opaque. Account removal policies can be governed by standards, auditing, and transparency requirements as in 4.1. This will be more effective in conjunction with user ID authentication, 1.2, so that a banned user cannot simply create a new account and resume egregious conduct.

4.6 *Account reputation scoring.* Account policing policies on social media can involve temporary demotion or other penalties short of outright bans. These instruments are also available for shaping user behaviors away from extreme and disruptive conduct, in accordance with 4.5.

4.7 *Microtargeting and narrowcasting.* As discussed 1.4, tailoring of advertising and content is a recent development of the information age enabled by large scale data collection about viewers and users, algorithms such as recommender systems, computing horsepower, and communications bandwidth. The policy options for addressing microtargeting by Malevolent Agitators apply as well to contemporary incarnations of broadcasting to the grassroots audience, and to internet delivery of targeted content by social media platforms that feature user-generated content.

4.8 *Transparency in sponsorship and advertising.* Individual ordinary citizen users on social media sites are able to attract sizable audiences that are subject to lucrative targeted advertising. Therefore, social media platforms have developed business models whereby content creators are paid for eyeballs and clicks. This is the intermediate zone (the bottom side of the Triad) between ordinary citizens and thought leaders possessing a platform and audience of fans or followers. The financial incentives that influence both content creators and the ads shown their audiences are shaping factors in public discourse.

4.8.1 *Ad transparency.* Require transparency in advertising and sponsorship for bloggers, podcasters, and influencers. Reportable information includes:
- o how content creators are paid, i.e. sponsorship, viewer profiles, watch metrics
- o how viewers are assigned ads to be shown with the content
- o identities of ad purchasers

4.9 *Digital Platform Agency.* Former FCC chairman Tom Wheeler and colleagues have proposed a Digital Platform Agency built around three concepts:

1. Oversight of digital platform market activity on the basis of risk management rather than micromanagement; this means targeted remedies focused on market outcomes and thereby avoiding rigid utility-style regulation,
2. Restoration of common law principles of a duty of care and a duty to deal as the underpinning of DPA authority, and
3. Delivery of these results via an agency that works with the platform companies to develop enforceable behavioral codes while retaining the authority to act independently should that become necessary.

("New Digital Realities; New Oversight Solutions in the U.S.: The Case for a Digital Platform Agency and a New Approach to Regulatory Oversight", Shorenstein Center on Media, Politics and Public Policy, August, 2020)

4.10 *Civil information defense and Cognitive Security.* "Cognitive Security" refers to the resiliency of a society against disruptors and unhealthy distortions of collective thought and decision making. Civil information defense (1.5) contributes to Cognitive Security. Other approaches and standpoints are warranted as well including mental health, economic resilience, and national cyberspace security.

4.11 *Local news media.* Local newspapers, radio stations, and other media have been decimated by the transition to the digital economy. This trend distances citizens from the government levels that they have most direct participatory and influence access to. Also, the trend elevates attention to national issues in ways that exacerbate nation-wide cleavages in thought and opinion.

4.11.1 *Financial support.* Legislation offering modest financial support for local news includes the "Saving Local News Act" and the "Local News Sustainability Act". Bolder measures to underwrite local news have been employed in Europe.

4.11.2 *Safe harbor.* A Safe Harbor exemption from antitrust concerns would enable local news organizations to bargain collectively with media platforms that present locally-generated content.

4.12 *Citizen participation in government.*

4.12.1 *Election integrity.* Election integrity is essential to citizen trust in government. This topic requires great sensitivity because it has become associated with partisan favor.

4.12.2 *Participation.* Citizens gain pride and trust in government when they participate and are recognized for doing so.

4.12.3 *Public service programs.* Publically funded programs like WPA, the Peace Corps, and Teach for America not only delivered services, but if run properly, they instill a tradition of community spiritedness and positive engagement with a diverse array of fellow citizens.

5 Adjust the Environment for Political Leaders
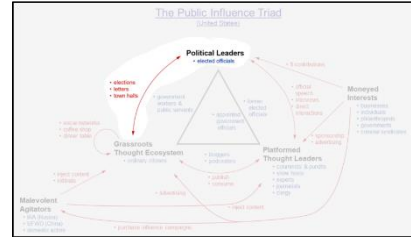
5.1 *Eliminate safe districts*. In many electoral races, Political Leaders can be successful by appealing exclusively to one side or the other of the political spectrum and dismissing the other.

5.1.1 *Gerrymandering.* Reform of electoral districting processes can create more balanced electorates that render extremist positions and rhetoric untenable.

5.1.2 *Voting methods.* Alternative voting methods such as Ranked Choice Voting encourage Political Leaders to cultivate moderate constituencies.

5.2 *Social conciliation*. When Political Leaders socialize together, they tend to be less acrimonious in their conduct about policy disagreements and less inclined to demonize the other side in their rhetoric. Institutional measures are available to encourage this practice.

5.3. *Cost of maintaining elected office.* Political Leaders' attention is distorted by the need to raise money to pay for election campaigns. Fundraising from Moneyed Interests incentivizes aligning with these interests' policy goals. Fundraising from the Grassroots incentivizes amplification of emotional appeals to identity and pushing hot button topics, to the detriment of measured debates about policy. See 2.2.