

How Social Media Fuels Populism and Strategies for Counteraction

Sejin Oh, Daeho Lee, Chehwan Kim

Hankuk Academy of Foreign Studies; sejin.sarah@gmail.com

ABSTRACT: This study aims to understand how social media exacerbates populist ideas and explain the implications of social media in the spread of populism, as opposed to traditional media forms. To comprehensively apprehend populist rhetoric and the role of the social media platform, this study conducts three experiments. An Echo Chamber Simulation experiment found that algorithmically reinforced discussion networks significantly increased exposure to ideologically homogeneous content and intensified agreement with populist framing. A Misinformation Susceptibility Survey revealed that simplified anti-elite narratives and misinformation led to an increase in the shareability of such headlines. In addition, a Populist Rhetoric Content Coding Project compared social media and traditional media and found that social media had more indicators of populist content. This study also includes a global policy case study of past policies to tackle Populism, such as the French Avia Law and the EU's Digital Services Act (DSA), comparing regulatory approaches aimed at limiting harmful online content while preserving democratic norms. The findings suggest that algorithmic amplification and network homogeneity play a central role in facilitating the spread of populist rhetoric online. Moreover, the study will conclude with policy recommendations that emphasize platform transparency, algorithmic accountability, and targeted moderation mechanisms designed to mitigate the spread of manipulative populist content without undermining freedom of expression.

KEYWORDS: Political Science, Political Communication, Populism, Social Media, Algorithmic Amplification.

■ Introduction

Division:

On January 6, 2021, thousands stormed the United States Capitol in an act of political violence that exposed the structural vulnerabilities of modern democratic governance. The insurrection culminated from a months-long disinformation campaign sustained on Facebook, amplified through Twitter threads, and coordinated across Telegram channels. A single lie—that the presidential election had been stolen—acquired mass belief through algorithmic amplification, exploiting the structural incentives of social media platforms designed to maximize emotional engagement. Each share, each reaction, and each repost fed the machinery, transforming a politically motivated fabrication into a widely held conviction that ultimately motivated real-world violence. This case illustrates a broader and increasingly global phenomenon: the systematic amplification of populist disinformation through digital infrastructure.

Comparable dynamics have unfolded across multiple national contexts. In Brazil, WhatsApp-disseminated disinformation shaped electoral decision-making at scale. In India, Facebook posts precipitated communal violence. Across Europe, far-right populist parties achieved unprecedented electoral gains through coordinated social media mobilization. These episodes, despite their geographic and cultural differences, share a common structural feature: populist sentiment intersected with platform architectures specifically designed to reward it.

How Populism Harms Democracy:

Populism, as a political logic, partitions the social world into two irreconcilable categories: a virtuous, homogenous “people” and a conspiratorial, self-serving “elite.” It operates through categorical moral claims that position one group in fundamental opposition to another, bypassing nuanced policy argumentation entirely. This binary logic repeats across diverse political contexts: Viktor Orbán frames Hungary against Brussels; American political figures invoke “ordinary people” against an unspecified establishment; Latin American strongmen present themselves as defenders of the nation against exploitation. The targets shift according to context, but the underlying formula remains constant.

What is historically novel is the speed and scale at which populism now operates. Populist ideas have always competed for political attention, but they have never previously possessed the capacity to reach millions instantaneously, to identify susceptible audiences through data analytics, or to coordinate those audiences without institutional intermediaries. Social media gave populism a structural political force of a fundamentally different order, one whose full implications scholars of political communication are still mapping.²

Digital populism disrupts three interconnected democratic functions. First, it eliminates traditional gatekeeping: the professional and institutional filters that once determined which political claims reached mass audiences have been circumvented, enabling misinformation to reach millions before factual correction reaches hundreds. Second, it fragments shared reality: when citizens inhabit algorithmically personalized information environments, they lose access to common factu-

al ground, and deliberative democracy—which depends on a shared factual basis for disagreement about policy—becomes structurally impossible. Third, it weaponizes online speech: political content no longer evaporates after circulation but functions as a coordination infrastructure, with a well-documented pathway from online belief to offline violence.³

Existing Literature and This Study's Contribution:

A substantial body of scholarship has examined the relationship between social media and political polarization. Pariser (2011) first theorized the “filter bubble” through which algorithmic personalization narrows users’ informational exposure, while Sunstein (2017) extended this analysis to argue that digital echo chambers systematically undermine the conditions for democratic deliberation. Research on populist communication specifically has documented the structural affinity between populist rhetoric—its emotional intensity, binary framing, and tribal identity appeals—and the engagement-maximizing logic of platform algorithms (Engesser *et al.*, 2017). Work on individual political contexts has examined how TikTok has become a primary vehicle for European populist radical right parties (González-Aguilar *et al.*, 2023) and how Twitter amplifies populist messaging among Portuguese and Italian political actors (Prior *et al.*, 2024). This study builds directly on these lines of research.

This study extends existing scholarship in three principal ways. First, whereas most prior research draws on observational data or secondary analysis, this study combines original empirical experiments with global policy case study analysis, producing a more integrated account of both the mechanism and potential regulatory responses. Second, this study treats social media’s relationship to populism as a structural property of platform architecture—the advantage populism enjoys on social media is by design, embedded in how these systems are built and incentivized. Third, this study incorporates a direct comparative analysis of traditional media and social media, documenting that social media amplifies populism more intensively and tracing the specific mechanisms responsible for that difference.⁴

Study Design and Structure:

To substantiate these claims empirically, this study conducted three original experiments. The first, an Echo Chamber Simulation, tested whether YouTube Shorts’ recommendation algorithm systematically pushes users toward politically extreme content as a function of their initial engagement patterns, using three anonymized accounts exposed to neutral, moderate conservative, and right-wing populist content, respectively. The second, a Misinformation Susceptibility Survey, examined whether populist-style rhetorical framing—characterized by emotional language, anti-elite appeals, and oversimplification—increases the perceived credibility and shareability of news headlines among Korean high school students, relative to neutrally framed versions of identical factual claims. The third, a Populist Rhetoric Content Coding Project, systematically analyzed political content across traditional media outlets (Fox News, CNN, BBC) and social media platforms (TikTok,

X, Instagram) during the 2024–2025 period, coding for the presence and intensity of standard populist indicators including anti-elite language, people-versus-elite framing, nationalist appeals, and emotional tone.

The paper proceeds as follows. Section II presents the three experiments in detail, reporting methods, results, and interpretations. Section III analyzes two major policy interventions—France’s Avia Law and the European Union’s Digital Services Act—as case studies in the regulatory challenge of combating digital populism without suppressing legitimate political speech. Section IV proposes three original policy recommendations derived from the experimental findings and policy analysis: algorithmic transparency mandates, friction-based sharing delays for populist-coded content, and the institutionalization of media literacy education. Section V synthesizes the study’s findings, situates them within the broader stakes for democratic governance, and identifies directions for future research.

Historical Timeline:

The relationship between social media and political populism has evolved through several distinct phases, each representing a qualitative transformation in how digital platforms interact with democratic processes.

2008: Barack Obama’s presidential campaign demonstrated that social media could mobilize voters at unprecedented scale. Contemporary observers characterized this development as ‘democratizing’.

2016: The Trump campaign and the Brexit referendum revealed that misinformation could propagate faster than factual correction, and that foreign actors could deliberately weaponize platforms for political interference.

2018: Brazil’s presidential election demonstrated that WhatsApp-disseminated misinformation could determine electoral outcomes. In India, Facebook posts preceded large-scale communal violence.

2020: The COVID-19 pandemic created an information vacuum that misinformation filled; millions rejected scientifically validated vaccines based on social media content, with preventable mortality as a consequence.

2021: January 6th closed the causal loop. Online conspiracy theory became physical insurrection, making the connection between digital and physical political action no longer theoretical.

2022–2024: Populist movements became structurally native to social media platforms. The infrastructure and the movements are no longer separable; each has shaped the other.⁵

This paper operates on one central thesis that social media has transformed populism from a political tendency into a structural political force. Four mechanisms explain this transformation. First, platform architecture rewards the emotional intensity that populism naturally generates. Second, algorithmic personalization sorts users into ideologically homogenous communities, reinforcing and intensifying pre-existing dispositions. Third, these platforms provide low-friction pathways from online sentiment to offline coordination. Fourth, populist disinformation travels through this system at speeds that

systematically outpace institutional correction. This matters because this distinction between ideas and infrastructure suggests an easier, more systematic response to defeating such an appeal, not one limited to rhetoric. At various points, it asserts that to stop digital populism, one must fight back against the intentional framework that accommodates and renders populist rhetoric easier.⁶

■ Experiment and Results

Introduction:

To comprehensively understand how populist narratives are spread by social media algorithms, we conducted three experiments. The first experiment focused on how social media algorithms, unlike traditional media, polarized users toward more extreme political content. The second was a survey that identified how susceptible to misinformation high school students were, especially for populist-style framing, while the third was a content coding project to measure how frequently populist rhetoric appeared in current media content. With Experiments 1 and 2, we attempted to measure how social media algorithms and populist-style framing influenced users' decisions, and finally, with Experiment 3, we compared and contrasted the rhetoric used in traditional media and social media for a complete narrative on populism.

Experiment 1. Echo Chamber Simulation:

a. Purpose

The purpose of this experiment was to test whether social media recommendation algorithms push users toward more extreme or homogenous political content, and to compare how initial viewing behavior, whether neutral, moderate conservative, or populist right-wing, affects later recommended content, and to provide support for the claim that social media is more prone to algorithmic polarization than traditional, non-personalized media.

b. Research Question and Hypotheses

This experiment aimed to answer two main research questions:

1. How does the YouTube Shorts recommendation algorithm change the political composition of a user's feed after a short period of targeted engagement?

2. Does prior exposure to populist/right-wing content lead to a higher proportion of populist recommendations compared to neutral or moderate conservative exposure?

This experiment had three main hypotheses.

1. Accounts that were fed populist/right-wing content will receive a significantly higher share of populist recommendations in subsequent Shorts.

2. Accounts that were fed moderate conservative content will show a noticeable drift toward more partisan and populist content.

3. Even neutral accounts will experience some degree of politicization due to algorithmic nudging, despite only being fed neutral content.

c. Experimental Design Overview

The platform used was YouTube Shorts, which is an algorithmically driven short-form video platform widely used by youth, with a strong "For You" and recommendation structure suitable for observing algorithmic drift. The experiment was conducted with three newly created and anonymous YouTube accounts. Account A was exposed to political news from YouTube channels without bias. Examples of these channels are Yonhap News and YTN. Account B was exposed to moderate conservative news, meaning that this channel's algorithm was influenced by right-wing media, but not those that are extremist. Some of them are TV Chosun and Channel A. Account C played videos classified as populist and right-wing populist content. For this account, content from personal channels was played. Each account was exposed to one type of political content during the "feeding" phase. For the time frame, each account had a single-session simulation using controlled viewing of 100 Shorts during the feeding phase and 100 shorts during the measurement phase per account, which isolates algorithmic reaction to the initial engagement in a short, controlled period. In total, the variables recorded are: the number of total shorts watched (100) and the number of political, neutral, moderate conservative, and populist shorts watched. Populist shorts were measured with the presence of certain keywords such as: 'vested interests,' 'political cartel,' 'national collapse,' 'media manipulation,' 'refugee crime,' 'media cartel,' and 'pro-North Korea forces'.

d. Results

The main results are summarized in Table 1 and Figure 1.

Table 1: A summary of the results for experiment 1. The results show a clear monotonic increase in populist content exposure from Account A to Account C, indicating strong algorithmic reinforcement of initial ideological signals. Even neutral engagement resulted in non-trivial political and populist recommendations.

Type of Short	Account A	Account B	Account C
Political	54	74	87
Neutral	43	4	4
Moderate Conservative	8	47	13
Populist	3	32	70

Account A, Account B and Account C

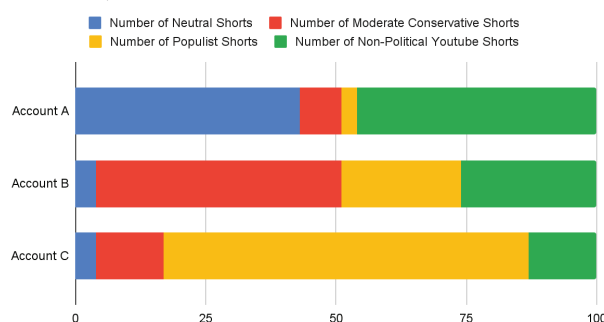


Figure 1: Graph of findings for experiment 1. The figure illustrates rapid algorithmic polarization, with populist-seeded accounts becoming overwhelmingly saturated with populist content and neutral accounts still receiving partisan recommendations. This demonstrates algorithm-driven ideological drift over a short exposure period.

Some immediate patterns become apparent, with a clear gradient from A → B → C becoming progressively more polarized. Account C, which is the populist account, becomes overwhelmingly saturated with populist content, and even the neutral account still receives non-neutral recommendations.

e. Interpretation

There are mainly three possible interpretations from this study. First is the algorithmic drift; even neutral signals produce political nudging by the algorithm, and moderate conservative signals intensify into significant populist exposure. Furthermore, populist signals create near-total echo chamber conditions. Second, algorithms escalate initial preferences through the echo chamber mechanism. The feed composition shifted toward extremity within 100 recommendations, which is at an astonishing speed. In contrast with traditional media, which present fixed content, not personalized feeds, the experiment demonstrates that personalization itself is the engine of polarization. Thus, this experiment provides support for Section II and suggests that social media is more dangerous than traditional media. This forms the basis for Section IV’s policy recommendations on algorithmic transparency and regulation.

f. Limitations

The experiment uses three YouTube accounts to test the hypotheses. Single-account-per-condition design limits generalizability, and broader replication would be needed for statistical confirmation.

Experiment 2. Misinformation Susceptibility Survey:

a. Purpose

The purpose of this example was to examine whether populist-style framing, including emotion, anti-elite language, or oversimplification, increases the public’s likelihood of believing or sharing content, regardless of whether the source is social or traditional media. For example, an example of a non-populist headline used in this experiment would be “Statistics Korea announced that household energy costs rose slightly this quarter,” as opposed to “Wages remain the same, but energy bills continue to rise, causing further financial hardship for households,” which is an example of a populist headline. This experiment critiques media messaging strategies as a whole and compares how populist rhetoric functions across all media forms, including both traditional media headlines and social media-style headlines, as this survey only analyzed populist dialogue, not the means of its spreading; thus, this experiment is a direct comparison of rhetorical effects, not platform effects.

b. Experimental Design

The survey consisted of 10 headlines in total, with four non-populist and neutral news headlines, four “populist-framed” versions of the same stories, and two misleading or false headlines. Participants were randomly divided into two groups: the control group viewed four neutral headlines and two misleading headlines, while the experimental group viewed four “populist-framed” headlines and the same two misleading headlines. Participants then rated each headline on

“Believability” and “Likelihood of Sharing,” each on a scale of 1 to 5. The “populist-framed” headlines only used rhetorical style manipulation. In this experiment, the independent variables were the three headline types (neutral/populist/misinformation), and the dependent variables were “Believability” (1-5) and “Willingness to Share” (1-5).

c. Research Question and Hypothesis

We hoped to answer whether populist-style framing increases believability and willingness to share through this experiment, and we hypothesize that populist-framed headlines will score higher on both believability and shareability. We also hypothesized that participants with stronger media literacy would be less susceptible to a populist framework.

d. Materials

First, our four neutral and populist-framed headlines were constructed with the same factual base but with different rhetoric. The eight headlines used are shown below in Table 2.

Table 2: Neutral and populist-framed headlines used in Experiment 2. Populist-framed headlines systematically employ emotional language, anti-elite framing, and oversimplification while preserving the same factual content as neutral headlines. This isolates rhetorical style as the primary independent variable.

	Neutral Headlines	Populist-Framed Headlines
1	Statistics Korea announced a slight increase in household energy costs this quarter.	Wages remain flat, but energy bills continue to rise, causing further financial hardship for households.
2	The Ministry of Education announced revisions to the standardized testing schedule for next year.	Education officials are unilaterally pushing through new testing regulations without consulting students and teachers.
3	City health authorities reported a rise in respiratory infections and recommended basic preventive measures.	Amid criticism that authorities are attempting to downplay the spread of the disease, citizens have no choice but to protect themselves.
4	Personal information protection authorities have begun reviewing a major tech company’s handling of user data.	Allegations have been raised that regulators have effectively turned a blind eye to the covert exploitation of user data by powerful tech companies.

Compared to the very neutral and non-extreme wording in the neutral headlines, the populist-framed headlines utilize populist rhetoric such as emotional language shown in “financial hardships on households”, negative portrayal of “officials” or “authorities,” anti-elite language such as “exploitation of user data by powerful tech companies,” and oversimplification throughout. This survey also presents headlines without source logos to eliminate brand bias.

The two misinformation headlines were:

- i. A recent study found that social media use increases the brain size of teenagers by 40% in just one year.
- ii. A new survey found that 90% of citizens plan to stop voting in the future.

These two headlines are obvious examples of false news and were used to measure participants’ media literacy.

e. Participants

The participants are Korean high school students, and thus, while this experiment may not reflect the entire population,

because adolescents are often high-engagement users of social media and will be exposed to more populist content, they are high-risk users. Thus, high school students are theoretically relevant to this survey.

f. Procedure

The survey was administered online via Google Forms, and the headline order was randomized to prevent anchoring. Also, the headlines themselves were generated via artificial intelligence to render completely neutral headlines. Participants rated each headline on Believability and Likelihood of Sharing, each on a scale of 1-5.

The participants are Korean high school students, and thus, while this experiment may not reflect the entire population, because adolescents are often high-engagement users of social media and will be exposed to more populist content, they are high-risk users. Thus, high school students are theoretically relevant to this survey. The total sample size is 59.

g. Results

The key findings are summarized in Table 3.

Table 3: Summary of findings for experiment 2. Neutral headlines outperformed populist-framed and false headlines on both believability and shareability. This suggests that populist rhetoric alone does not increase persuasion absent algorithmic amplification.

1. Mean believability for neutral headlines (neutral_b)	3.62
2. Mean believability for populist headlines (populist_b)	2.79
3. Mean believability for false news headlines (false news_b)	2.11
4. Mean shareability for neutral headlines (neutral_s)	2.95
5. Mean shareability for populist headlines (populist_s)	2.73
6. Mean shareability for false news headlines (false news_s)	2.37

Neutral headlines scored higher for both believability and shareability by 0.83 and 0.22, respectively; the difference between neutral and populist headlines by shareability is notably smaller than in believability. Across all measures, neutral headlines were rated more credible and more shareable than populist or false news headlines.

h. Discussion and Interpretation

The pattern was that in both believability and likelihood of sharing, neutral headlines scored the highest, followed by populist and false news headlines. This means that while populist framing still increases believability relative to misinformation, it underperforms neutral headlines, meaning students are not blindly reactive to emotional or oversimplified cues, suggesting susceptibility to populism is not absolute. Thus, this experiment empirically shows that framing and rhetoric drive perception, and shows that populist rhetoric can boost misinformation-like reactions even without false content, which should be accounted for in policy contexts. Given these findings, it is essential to examine how this dynamic has been acted out in real-world political crises and policy environments.

The most interesting takeaway from this experiment was its deviation from the hypothesis. We hypothesized that populist headlines would outperform neutral ones in both believability and shareability because of their provocative nature. Instead, we observed that neutral headlines outperformed both. Therefore, if populist framing is not inherently more persuasive absent algorithmic amplification, it reinforces the core argument that platform architecture, not rhetoric alone, is the key variable. Although the results were the opposite of the hypothesis, this result in turn strongly supports our main claim.

The pattern was that in both believability and likelihood of sharing, neutral headlines scored the highest, followed by populist and false news headlines. This means that while populist framing still increases believability relative to misinformation, it underperforms neutral headlines, meaning students are not blindly reactive to emotional or oversimplified cues, suggesting susceptibility to populism is not absolute. Thus, this experiment empirically shows that framing and rhetoric drive perception, and shows that populist rhetoric can boost misinformation-like reactions even without false content, which should be accounted for in policy contexts. Given these findings, it is essential to examine how this dynamic has been acted out in real-world political crises and policy environments.

Experiment 3. Populist Rhetoric Content Coding Project:

a. Purpose

The purpose of this experiment was to measure how frequently populist rhetoric appeared in current media content and to compare the rhetoric used in traditional media vs. social media. It analyzes both traditional media, such as newspaper opinion pieces and TV news clips, and social media such as TikTok, X (Twitter), and Instagram, allowing for direct cross-media comparison. Moreover, this experiment supports Sections I and II by operationalizing what populism looks like in the real world and provides empirical grounding for later policy recommendations.

b. Methodology and Data Collection Design

First, the research period was between November 2024 and November 2025, focusing on the 2024 election period and recent developments. To compare two distinct media sources, this experiment analyzed Fox News, CNN, and BBC for traditional media sources and TikTok, X(Twitter), and Instagram for social media platforms. Geographically, this experiment focused on Western countries, including the United States, the United Kingdom, and European countries such as Portugal, Spain, the Netherlands, Finland, Germany, France, and Italy.

c. Operationalization

Indicators used in the experiment to measure how biased or populist an article was include:

- Anti-elite language
- People vs. elite framing
- Nationalist purity
- Conspiracy framing
- Anti-immigrant appeals
- Emotional tone (low/medium/high)

- Aggression and Polarization

The coding sheet structure included the headline and a snippet of the content, the source and date, topic category, and whether there were populist markers present, marked by yes/no, and the intensity. Furthermore, the rhetorical style was also noted, and some examples of populist rhetorical style markers are: "crisis framing" and "enemy construction".

d. Analytical Strategy

First, this experiment compared the frequency by counting the presence of each indicator per medium and normalizing by sample size. Moreover, because traditional media tends toward indirect populism, for example, by quoting politicians, and social media tends toward direct and emotional populism amplified by algorithms, this was also taken into account. The strategy used in this experiment also included topic-level contrast, and such topics included: immigration, corruption, economy, national identity, and quantified differences in rhetorical temperature.

e. Results

i. Traditional Media Analysis

Fox News operates through a "counter-elite news brand" that fuses populist political rhetoric with tabloid journalism style, hailing its audience as "the real Americans" while representing conservative political demands as popular and universal. For example, Fox News has coverage framing Marine Le Pen's conviction as "European governments lashing out at populist leaders," with Trump commenting that her situation "sounds like this country". This reflects a broader shift in Fox News' atmosphere, shifting from being skeptical of Trump to a "Trump safe space, with a dose of Bannionist populism," with academic studies finding that Fox's programming engages in rhetorical themes similar to propaganda. Republican columnist Jonah Goldberg acknowledged Fox "does lean to the right" and described it as "less a bastion of ideological conservatism and more a populist, tabloid-like network". As for populist indicators, Fox News had examples of anti-elite language when it consistently framed mainstream institutions as corrupt, people vs. elite framing with their "real Americans" narrative, conspiracy framing with deep-state narratives, Nationalist appeals with their widespread "America First" rhetoric, and anti-immigrant appeals with border crisis coverage.

Second, this experiment analyzed CNN's coverage from the 2024 election, which focused primarily on factual reporting of electoral results, Trump's victory speech, intelligence warnings about foreign interference, and voter testimonials. Post-election CNN coverage included Latino voter perspectives, with Juan Sanchez stating "Democrats just offer, they don't deliver" and describing the government as "a disaster," while Roberto Mendoza felt "positive changes are coming". For populist indicators, there was limited direct populist framing from CNN itself, with voter quotes containing anti-establishment sentiment and more neutral and procedural election coverage. There was also a less emotional and aggressive tone compared to Fox News.

Third, we analyzed BBC's coverage on immigration in November 2024, when BBC reported on the UK government's asylum system overhaul, with Interior Minister Shabana Mahmood stating undocumented migration is "tearing the country apart," while announcing plans to cut refugee status from five years to 30 months. BBC coverage quoted Mahmood saying, "People can see huge pressure in their communities". For populist indicators, there was Anti-elite messaging, although indirect and through government sources, and there was an overall focus on a system failure narrative. However, there was only a moderate nationalist framing, and overall, the BBC emphasized control over immigration.

ii. Social Media Analysis

First, an analysis of 51,680 political TikTok videos from the 2024 U.S. presidential cycle revealed that toxic and partisan content consistently attracts more user engagement, with 77% of videos being explicitly partisan and associated with approximately twice the engagement of nonpartisan content.⁷ Populist Right Parties have used this to garner widespread support. For example, European populist radical right parties like Spain's Vox and Germany's AfD effectively use TikTok to disseminate anti-elitist and anti-immigration content while exploiting platform features through entertaining and inspiring young audiences, with identity-based "us versus them" narratives targeting migrants or "woke leftists" as enemies. Specifically, Nigel Farage of the UK Reform Party's TikTok campaign addressed immigration in only 7.7% of videos, but used right-wing populist tactics to create "Us" and "Them" binary distinctions premised on xenophobia, stating "If you're wondering why you can't get a house, why your rent is up 20% over the last three years, if you're wondering why your granny can't get a GP appointment, it's because we've had a population explosion." Another example is the Finnish Finns Party, which presented itself as representing "the people" against elites, with party leader Riikka Purra portraying other parties as "out of touch with people's everyday concerns, focusing on climate change reforms while they neglect people's access to basic services," while framing opposition to immigration as "offering solutions" rather than being driven by racism. To summarize with an analysis of populist indicators on TikTok, there was very frequent anti-elite messaging utilizing a framing of people vs. elites as their core strategy. The algorithm then amplified such indicatives in an emotional and aggressive tone, and appealed to the extreme right through anti-immigrant appeals as their central theme. There was also an overall nationalist purity through identity-based messaging. This can be seen in the fact that videos about immigration and election fraud topics draw high levels of toxicity and attention, with toxic videos on immigration receiving 3.5% more engagement.

Second, this experiment analyzed Portuguese Populist Leader André Ventura of the Chega Party's posts on X. An analysis of Ventura's Twitter posts during the 2022 election revealed a compact agenda where "democratic regeneration" (21%) and "corruption" (39.1% combined with related themes) dominated his messages, presenting himself as the legitimate representative of "common people" against "corrupt elites who

for decades have pursued partial interests." Ventura strategically uses news to promote themes like nationalism, immigration control, corruption, and social dichotomy between "us" and "them," leveraging news headlines as heuristic shortcuts to reinforce political messages. This reflects some broader patterns on X. For example, an analysis of 1,115 tweets from Portuguese political leaders during 2024 campaign showed populist far-right leaders privilege subjects such as "Insecurity," "Justice, corruption and clientelism," with right-wing politicians using predominantly right-wing discourse while populists employ typical ideological resources like "emphasizing sovereignty of the people, advocating for the people, attacking the elite, ostracizing others". Furthermore, a study of 390,000 posts by over 700 Italian politicians on Twitter-X found that populist rhetoric indeed increases the likelihood of posts going viral, even among non-populist voters, with the opposition between corrupt elites and virtuous people aligning well with the simplification and polarization that guides social media storytelling. Trump's posts on X also reflect the broader atmosphere of the platform. For example, Trump's promises to "drain the swamp" and "demolish the deep state" remained central to his 2024 campaign messaging, with a ten-point plan released to "shatter the deep state" that includes removing "rogue bureaucrats," cracking down on government "leakers," and creating term limits for Congress. To summarize with populist indicators on X, there is extremely frequent anti-elite messaging with corruption framing as a dominant theme, conspiracy narratives such as the deep state and the "system" being rigged, a core binary narrative of "People vs. elites," and a strategic use of anti-immigrant messaging.⁸

iii. Comparative Analysis of Traditional vs. Social Media

A summary of the key findings of the frequency of populist language is summarized below in Table 4.

Table 4: Comparison of the frequency of populist content in traditional and social media. Social media exhibits substantially higher emotional intensity, direct attacks, and people-versus-elite framing than traditional media. This highlights the role of platform affordances in amplifying populist rhetoric.

Indicator	Traditional Media	Social Media
Anti-Elite Rhetoric	Moderate (Fox) Low (CNN/BBC)	Very High
People vs. Elites	Present (Fox) Limited (CNN/BBC)	Central Theme
Emotional Intensity	Moderate	Very High
Conspiracy Framing	Present (Fox) Limited (CNN/BBC)	Frequent
Anti-Immigration	Regular Coverage	Dominant Theme
Direct Attacks	Indirect/Quoted	Direct/Personal

Social media toxic videos were associated with 2.3% more interactions overall, with election-related toxic content receiving 1.3% more interactions and immigration-related toxic content receiving 3.5% more engagement.

A key difference between traditional media and social media is that, while traditional media is more filtered and indirect, except for Fox News, social media is raw and unfiltered, and

populism is algorithmically amplified.⁹ Authoritarian populism combines features where leaders rhetorically divide the population into majority versus elites while positioning themselves as true representatives, using fearmongering and scapegoating marginalized groups, and this is particularly apparent in social media. This experiment also examined platform-specific amplification, and research found TikTok's recommendation algorithm exhibited pro-Republican bias during the 2024 presidential election, with Republican-seeded accounts receiving approximately 11.8% more party-aligned recommendations, while Democratic-seeded accounts were exposed to approximately 7.5% more opposite-party recommendations, and videos on immigration, crime, and foreign policy were most likely to be recommended as ideological mismatches to Democratic-conditioned accounts.

iv. Topic Analysis

There were four most common populist topics in total.¹⁰ First is immigration and border control. While traditional media is more policy-focused and statistics-driven, social media is more driven by emotions and frames this topic as a crisis. For example, populist parties VOX (Spain) and Chega (Portugal) employ confrontational discourse with the political elite and emotional appeals to mobilize followers, with VOX attacking "the left" while Chega criticizes "the political system as a whole." Second is corruption and elite malfeasance. While traditional media focuses on investigative and evidence-based reporting, social media is often accusatory and adjacent to conspiracy theories, which was a dominant theme across all populist social media accounts. Third is the economy and class issues. Traditional media focuses on economic data and expert analysis, while social media is more grievance-based and focused on blame. Placed in a real-world context, economic stress and cultural disruption drive political shifts, with financial crises consistently resulting in significant gains for far-right and left-wing populist parties across 800 elections in 20 democracies from the 1870s to 2014. Fourth is national identity and sovereignty, and while traditional media focuses on the historical context and policy implications for this topic, social media appeals emotionally and is fear-based, appealing to "heartland" and "real" citizens.

f. Extended Analysis

The last part of this experiment is a comparison of low-trust traditional media vs. social media in populist impact, specifically comparing media bias, trust ratings, and populist reception. This experiment analyzed media trust following the Ad Fontes and YouGov framework, which media sources on two scales: political bias (left to right) on the horizontal axis and reliability on the vertical axis, with sources concentrated in an "inverted-U" shape where neutral bias sources are generally reliable in original fact reporting, while sources with extreme bias on either side often contain factually inaccurate information and propaganda.

The most trusted YouGov sources included The Weather Channel, which has a +49 net trust score, meaning Americans are 49 percentage points more likely to call it trustworthy

than untrustworthy. BBC (+26), PBS (+25), and the Wall Street Journal (+22) followed. The least trusted include: The National Enquirer (-37), Infowars (-15), and Breitbart News (-7). Moreover, analysis reveals that left-leaning outlets cluster around the 40-point reliability mark between "skews left" and "strong left" columns before dropping rapidly in credibility, while on the conservative side it's a much more linear trend, with WSJ more reliable and less biased than Fox, which is more reliable than Breitbart, down to Turning Point USA, down to InfoWars with Alex Jones. In a 2017 survey among US readers, Breitbart News was voted the third least trustworthy source, with an October 2018 Simmons Research survey ranking it the sixth least trusted news organization, while an August 2019 internal Facebook study found Breitbart News was the least trusted news source and ranked as low-quality across the U.S. and Great Britain. Under Bannon's management, Breitbart News aligned with the American alt-right, the European populist right, and the pan-European nationalist identitarian movement, with Bannon declaring the website "the platform for the alt-right" in 2016. By 2019, Breitbart had lost nearly 75% of its readership, going from 17.3 million monthly readers at the beginning of 2017 to 4.6 million in May 2019, and by October 2024, its monthly readership had declined to around 700,000.

Another important phenomenon to address is the "Trust Crisis." According to Pew Research Center surveys from September 2025, 56% of U.S. adults say they have at least some trust in information from national news organizations, down 11 percentage points since March 2025 and 20 points since 2016, while trust in local news organizations stands at 70%, also down from 80% in March and 82% in 2016. Trust in mainstream media has plummeted to just 30 percent among American adults in 2022, with Gallup polling in 2024 showing only 32 percent of Americans have a "great deal" or "fair amount" of trust in mass media, marking the third consecutive year of declining trust levels.¹¹ Interestingly, surveys show citizens may trust social media more, as demonstrated by the Social Media Trust Paradox. A 2024 PRWeek survey of 155 PR professionals found 50% say consumers trust social media the most by a wide margin, with traditional news media following at 39%, followed by business (10%) and government (1%). There is also a partisan trust divide; following assaults on mainstream media by GOP leaders, 53 percent of Republicans found mainstream news not at all trustworthy, compared to just seven percent of Democrats and 35 percent of independents, leading Republicans to trust personal stories from friends and family as much as articles sourced from multiple experts. Fewer than half of Republicans and Republican-leaning independents (44%) now say they have at least some trust in information from national news organizations.¹²

However, within traditional media sources, not all content on Fox News, MSNBC, or CNN is the same, with individual shows from CNN, Fox News, and MSNBC falling in different categories of bias and reliability, such as Fox News shows Outnumbered (lower right) and Special Report w/Bret Baier (top center) being very different according to Ad Fontes analysis.

Audiences also respond differently to populist rhetoric from low-trust sources. First, audiences show an "alternate media for populist audience" phenomenon. German research on alternative media with affinity to populism (AMP), including Breitbart and Infowars, revealed substantial differences between occasional and frequent users, with frequent AMP exposure related to higher personal relative deprivation, stronger populist attitudes, and higher likelihood to vote for right-wing populist party AfD, suggesting frequent AMP use can be interpreted as partisan selective exposure. This shows that users with existing populist attitudes seek out low-trust sources that confirm their worldview. Second, there is a credibility transfer from message to medium. Research on social media credibility found that the reliance on a social media platform to get news has the strongest impact on how credible that social media platform is perceived, with the more trust social media users have in alternative news sources, the more credible they perceive social media. The critical mechanism demonstrated here is that perceived credibility of a news source positively transfers from a traditional news source to an online modality, suggesting that source credibility carries over to the credibility of the online medium. Third is the "disinformation accusation" effect, and experimental research (N=1,330) showed that politicians' disinformation accusations reduce audience members' trust in the accused news outlet and perceived accuracy of news messages, while trust in the accusing politician remains largely unaffected, with only individuals with strong populist attitudes generalizing disinformation accusations to the media as an institution. This implies that low-trust sources can further damage mainstream media credibility among populist audiences without harming their own standing. Last, there is a spread of disinformation by source type. A German experimental study (N=572) showed that the spread of disinformation is mediated by perceived message credibility, with individuals with populist radical-right (PRR) attitudes particularly inclined to spread disinformation, though it was unclear whether this was due to the large amount of PRR-biased disinformation or a general tendency to perceive disinformation as credible.

Meanwhile, trust in news on social media platforms has increased over the past year, to an average of -10 across platforms from -16 in 2024, largely due to increasing trust among Republicans, with average net trust rising 15 points to -3 from -17 in 2024. Another phenomenon to take note of is the Trust-Use Paradox; for partisan outlets like France's CNews (regarded as French equivalent of Fox News), trust precedes use, illustrating trust's versatile role where selection effects guide individuals to choose attitude-congruent media sources, while for traditional outlets like BFMTV and TF1, news trust is an effect of media consumption built over repeated exposure. The Credibility Trivialization Effect is also important to take into account. For example, German research suggests that for individuals with populist radical-right attitudes, credibility becomes trivial when the message aligns with their beliefs, with evidence showing PRR bias of disinformation may explain the spread of disinformation by individuals with PRR attitudes despite a lack of credibility.¹³ This suggests that source trust matters less when the message confirms a pre-existing populist worldview.

This experiment also compared low-trust traditional media and social media in terms of populism. The mechanism comparison table is shown below in Table 5.

Table 5: Mechanism comparison table. Traditional media reaches mainly self-selected, reputation-conscious audiences, while social media spreads content more widely through algorithms and weaker source awareness.

Factor	Low-Trust Traditional Media	Social Media Platforms
Believability	Extremely Low	Moderate to Low
Audience	Strong ideological alignment required	Wide Audience (Algorithm-driven by users' choice)
Populist Message Receptivity	High among existing believers	Amplified by engagement metrics
Editorial Standards	Minimal editorial standards	None
Sharability	Limited by reputation damage	Enhanced by platform affordances
Perceived Believability	Low but stable	Variable

We hypothesize that social media amplifies populism more than low-trust traditional media due to a variety of reasons. First is algorithmic amplification as opposed to editorial gate-keeping. Even low-trust traditional media maintain some editorial process. Social media algorithms favor polarizing content that triggers strong reactions, with TikTok's affordances favoring amplification of right-wing content and different populist themes generating varying levels of user engagement.¹⁴ The second reason may be perceived authenticity; PR professionals identify "other consumers" as the top choice for whom consumers find "very trusted," suggesting social proof and peer-to-peer messaging carry more weight than institutional sources. Third, as opposed to low-trust traditional media in which audiences know they're consuming "Breitbart" or "InfoWars" with all attendant baggage, in social media, populist messages appear alongside trusted content, from friends, and without clear source attribution, which may make populist messages seem more trustworthy. Lastly, engagement metrics on social media reward populism. An analysis of 1,777 campaign tweets from 22 national-level actors across five countries found that populist discursive frames are associated with higher online engagement compared to alternative discursive frames, including pluralism, technocracy, and neutral discourse.

In conclusion, the Trust Paradox Model for Populism-Trust-Medium interaction for low-trust traditional media as opposed to social media:

1) LOW-TRUST TRADITIONAL MEDIA:

Known Low Credibility → Self-Selected Audience → Preaching to Choir → Limited Spread

2) SOCIAL MEDIA:

Variable/Unknown Credibility → Algorithm-Selected Exposure → Broader Reach → Viral Spread

However, it is important to mention that there are several shortcomings to this study. The sample could potentially be biased. For experiment 1, the study conducted was a single-run trial, so the results may not reflect an overall trend in other accounts or devices. There is also no way of knowing whether the algorithm was influenced by the previous watching habits of the previous users, as the devices used were not brand new. There is also no true causal isolation, as we cannot realistically prove the algorithm caused the shift rather than interaction effects between content type and engagement mechanics. There was also little control over confounding variables that might have affected the algorithm, including watch

time, scrolling speed, or micro-interactions like pauses and re-watches. Moreover, this experiment only used YouTube Shorts, so findings cannot be generalized to other platforms such as TikTok, Instagram Reels, or X, which have vastly different recommendation architectures. The "feeding phase" used in the experiment also has limitations in that it is only an artificial emulation of experimental conditions. Real users would have inconsistent, multi-topic, and long-term engagement histories, so this reduces ecological validity. There may also have been a bias in the researcher, as the coding of "populist content" depends on selected keywords or subjective classification, and researchers' biases may have labeled content as populist; this is exacerbated as there was no inter-coder reliability reported. For experiment 2, because of logistical limitations, the sample size is 59, and all research subjects are Korean high school students with high academic profiles; thus, this sample would be more likely to be one that would be better at spotting populist or biased rhetoric. Therefore, this sample could not reflect the sentiment of the general public. Furthermore, while the experiment isolated headlines to judge reactions based on language alone, because the headlines lack source attribution, real-world credibility cues are removed. This experiment may also be victim to the self-report bias, as 'believability' and 'likelihood of sharing' are stated preferences that do not equal actual behavior. For experiment 3, our research examines a limited set of media, comparing traditional media such as Fox News, CNN, BBC, versus social media such as TikTok, X, and Instagram, which is not representative of the full media ecosystem, which could potentially lead to cherry-picking bias.

Overall, there may be a theoretical overreach, as our research claims structural causation by stating that the platform architecture drives populism when empirical design mainly shows correlation and short-term amplification, not systemic causality. The experiments also do not account for the ideology, prior beliefs, or political interests of the subjects and researchers.

■ Discussion: Global Policy Case Studies

France's Avia Law:

France's Avia Law (2020) was one of Europe's earliest attempts to cut down on malicious online content aggressively. The law was created to delete terrorist posts and hate speech on major platforms and search engines, requiring sites to delete user-reported hateful content within 24 hours, and terrorist or child-abuse material within one hour. Failure to do so resulted in substantial fines (Freedom House). Although not directly mentioning "populism", the policy was clearly created to combat xenophobic, racist, and anti-minority content, which are often intertwined with right-wing populist actors. The aim was to quickly delete harmful content to prevent radicalized material from spreading.¹⁵

This reveals how social media fuels populism. Populist movements are characterized by their emotionally-charged, -based attacks targeting immigrants, religious minorities, or globalist elites, which easily become hate speech. The strict deadlines imposed by the Avia Law attempted to disrupt this vicious cycle, to cease the creation of openly racist, antisemitic, or Is-

lamophobic messages.¹⁶ Moreover, this gave platforms a strong sense of responsibility, identifying them as gatekeepers for appropriate content.

However, the French Constitutional Council ruled that there were elements of the law that were unconstitutional,¹⁷ owing to their extremely short deadlines and large fines. Besides, no prior judicial review raised a disproportionate infringement on freedom of expression (Freedom House). If the bill passes, platforms will have strong incentives to remove any content that could be viewed as risky, relying on strong filters and strict decisions.¹⁸ This poses two threats to democracy. Legitimate political speech could be eliminated, making the online space no longer a place for a fruitful debate. Additionally, this targets individual content, not the algorithmic systems that generate echo chambers, highlight sensational posts, and promote misinformation. Essentially, it misses the more effective and urgent solution.

After the ruling, softer measures were implemented to regulate harmful material. Methods included specialized online hate prosecutors and easier reporting methods. These surviving elements were better options than giving platforms an immense burden. The Avia Law, therefore, illustrates the necessity and the difficulty of regulating populist content. Aggressive laws may be tempting, but a drastic change undermines constitutional protections while neglecting the urgent problems created by the algorithm.

The European Union (EU)'s Digital Services Act (DSA):

While France's Avia Law attempted to delete individual content, the European Union's Digital Services Act (DSA) used a more strategic method to remove populist content. Enforced in 2022, it provided an extensive framework for censorship, giving specific responsibilities to "Very Large Online Platforms" and search engines, in order to target algorithms that create echo chambers (European Commission).

DSA represents identifying and mitigating broader risks, tackling the service as a whole. This prevented harm to fundamental rights, civic discourse, and electoral processes (European Commission). This helped adjust recommender systems not to promote misleading or inflammatory content, giving users less-manipulative options by utilizing fact checkers.¹⁹

Positioning itself as one of the first actions to cut down on the amplification of populist ideas, the system-level approach was able to modify recommendation engines to allow more responsible content to be produced.

Comparison:

Avia Law focuses on removing individual harmful posts through strict takedown deadlines and even imposes 1-hour and 24-hour deletion requirements for terrorist and hateful content, treating populist rhetoric as a post-by-post problem. It also forces platforms to act as rapid-response moderators, incentivizing maximal censorship to avoid fines. However, it was struck down for disproportionate restrictions and a lack of judicial review, endangering legitimate political speech. In summary, it disrupts surface-level harmful posts but fails to address the algorithmic forces that allow populist messages to

go viral, and illustrates the risks of aggressive deletion regimes such as constitutional vulnerability, over-censorship, and minimal impact on systemic misinformation. On the other hand, the DSA focuses on regulating platform systems such as algorithms, recommender engines, and structural amplification mechanisms. It also mandates comprehensive risk assessments and mitigation strategies for "systemic risks," treating populist spread as an ecosystem-level problem, forcing platforms to increase transparency, adjust recommender systems, cooperate with fact-checkers, and undergo independent audits. The DSA also preserves a wider expressive space by targeting amplification structures instead of directly suppressing user content, working to reduce the reach of misleading or inflammatory populist content by redesigning platform architecture rather than deleting individual posts. Thus, it establishes a model for sustainable regulation by shifting focus from content removal to structural governance of digital platforms.

From comparing these two models, we can see that the Avia Law may be overly aggressive, as speed-based content deletion creates constitutional risks, incentivizes over-censorship, and fails to address the deeper drivers of populist virality. From the DSA, we learn that structural, system-level regulation that works to target algorithms, recommender systems, and amplification mechanisms produces more sustainable moderation without suppressing legitimate political speech. Thus, we can interpret that an effective synthesis of the two policies to effectively counter populism requires shifting away from post-by-post policing toward regulating the digital architectures that reward sensationalism and misinformation. Democracies must balance rights protection with intervention by adopting transparent, proportionate, and system-oriented rules rather than punitive removal mandates and creating policies that combine Avia's recognition of harmful rhetoric with the DSA's structural approach to offer a pragmatic model that regulates amplification, protects debate, and avoids the pitfalls of rapid takedown regimes.

■ Discussion: Policy Recommendations

Our research has established a straightforward fact: social media platforms amplify populism not because populist arguments persuade more people, but because platform architecture rewards the specific characteristics populism naturally possesses. Emotional intensity. Dichotomous framing. Tribal belonging. The response must therefore address the mechanism. Banning populist speech will fail. It will only drive populism to alternative platforms. Convincing populists to abandon populist frameworks will fail. These frameworks address genuine grievances and emotional needs that alternative framings do not address. The only viable response is to alter the mechanism that gives populism a systematic advantage.

This means three specific interventions.

The first recommendation is to make algorithms visible. The first principle of sound governance is this: what you cannot see, you cannot govern. Democratic societies cannot govern what remains hidden. Therefore, algorithms must become visible.

Platforms publish quarterly reports detailing:

- What factors algorithms use to rank content (not how, but what)
- What percentage of each user's feed comes from algorithmic ranking versus explicit user choice
- Whether recommendation patterns differ by the political orientation of seed accounts
- Independent audits, conducted annually by accredited research institutions, must answer these questions:
- Do algorithms disproportionately amplify populist, conspiratorial, or inflammatory content?
- By what magnitude does algorithmic mediation push users toward ideological extremes?
- Do these patterns differ across platforms?

Why does this matter? Because without visibility, reform is impossible. Currently, platforms claim that algorithmic effects are beneficial or neutral, or trivial. No one can verify these claims. With visibility, competing claims can be measured against evidence. Some platforms will boast of diverse recommendations. Others will be forced to confront that they produce polarization. Reform will follow from transparency.²⁰

The second recommendation is to introduce friction without suppression. A paradox of human behavior is worth noting: we behave differently when we pause, however briefly, to consider what we are about to do. Facebook has discovered this through suicide prevention interventions. When users indicate suicidal intent, Facebook does not remove their posts. Instead, it inserts a pause—an opportunity to reconsider. This modest intervention, studies show, measurably reduces harm without suppressing speech.²¹

The same principle can apply to populist content.

Here is the mechanism:

When posts meet algorithmic criteria for populist-radical-right framing—people-versus-elites dichotomies, conspiracy narratives, inflammatory characterization of out-groups—they trigger a 15-second sharing delay. Users see the content. They can read it. They can click through to engage with it. But they cannot instantly amplify it to their networks. They must pause. They must consciously choose to share. This introduces friction while not actively suppressing speech. The user retains complete agency. People share without reading. They amplify without thinking. A brief pause interrupts the reflex without preventing deliberate action.

Why is this superior to censorship? Because it preserves user agency while altering the structural advantage populism currently possesses. Users who want populist content can access it. They can share it. They are simply not given an algorithmic boost for doing so. This is what platforms do for spam and low-quality content already. The question is not whether platforms can implement such systems but whether they will choose to.

The third recommendation is to build media literacy as an infrastructure. We know from our research that neutral information outperforms populist-framed information when receivers possess media literacy. We also know that media literacy is not evenly distributed. Working-class populations, elderly populations, and populations in information deserts possess a lower capacity to resist populist framing. Democra-

cies have always invested in basic literacy—the ability to read. Media literacy—the ability to understand how information systems work and how to evaluate claims—is equally foundational to modern democratic participation.²³

In secondary schools, a mandatory 20-hour curriculum must teach:

- How recommendation algorithms function and how they differ from human editorial judgment
- What populist rhetorical frameworks are, and what emotional appeals they employ
- How to recognize when emotional intensity is substituting for evidence
- How to consume information across the ideological spectrum without absorbing any single narrative uncritically
- How conspiracy theories differ from legitimate institutional critique

Public service media campaigns must explain:

- Why platforms amplify certain types of content
- What red flags indicate conspiracy-adjacent framing
- The distinction between legitimate criticism of institutions and populist false dichotomies²²
- Where to find reliable information sources across the political spectrum

Why does this matter? Because educated citizens are more resistant to manipulation, not because education makes people progressive or conservative, but because it teaches people to ask questions of claims rather than accept claims reflexively.

■ Conclusion

This study set out to answer a central question: why does social media amplify populism with a consistency and intensity that other media have not produced? The findings across three original experiments converge on a single structural explanation: platform architecture systematically rewards the specific characteristics that populism naturally possesses.

The Echo Chamber Simulation demonstrated this most directly. Initially neutral YouTube accounts drifted toward 54% political content within a single session; moderate conservative accounts reached 70% political content with 31% populist saturation; and populist-primed accounts received 87% political content with 70% populist representation. The gradient was proportional to initial ideological intensity, showing that algorithms amplify extremity through engagement-maximization logic that rewards emotional intensity and tribal belonging—the same properties that define populist content.

The Misinformation Susceptibility Survey introduced an important qualification. Populist framing, in isolation, proved less persuasive than neutral framing—believability scores of 2.79 versus 3.62, shareability of 2.73 versus 2.95. The amplification that makes populist content appear more persuasive in practice is a product of the algorithmic contexts in which it circulates, where structural exposure compounds individual susceptibility. Populism's rhetorical power and its algorithmic power are distinct and should be addressed as such.

The Populist Rhetoric Content Coding Project situated these findings in real-world media. Toxic content received 3.5% higher engagement on immigration topics; populist rhetoric increased viral probability on Twitter; and platforms exhibited systematic bias toward partisan content. Fox News engaged in populist framing consistently, while CNN and BBC maintained a substantially lower rhetorical temperature. Across social media—TikTok, Twitter, and Instagram—populist rhetoric, conspiracy framing, and anti-elite messaging were the dominant register, present at levels that have no equivalent in mainstream broadcast or print media.

Why This Matters for Democracy:

Democracy operates on a specific assumption: citizens disagree about policy but agree on facts. When this assumption holds, deliberation is possible. Citizens can debate what policies will best address agreed-upon problems. When this assumption breaks down—when citizens cannot agree on basic facts—deliberation becomes impossible. Social media fragments shared reality. Our research documented this process.

Algorithms customize information environments to individual preferences, pushing users progressively toward ideological extremes. Users begin with moderate exposure. Algorithmic mediation pushes them toward increasingly partisan, conspiratorial, and emotionally intense content.

The consequences for democracy are severe. If citizens cannot establish shared factual ground, democratic deliberation collapses. If populist narratives consistently reach larger audiences through algorithmic amplification, and institutional narratives reach smaller audiences, populist alternatives gain credibility relative to institutional responses. Education and rationality cannot overcome structural incentive systems. If platforms reward inflammatory content with engagement, users who produce inflammatory content will do so even when they recognize it is inflammatory.

What Structural Reform Requires:

Structural reform requires altering the machinery that gives populism a systematic advantage while preserving democratic pluralism. Three interventions are necessary and sufficient: making algorithms transparent so citizens understand the forces shaping their information environment; introducing friction into algorithmic amplification without suppressing speech; and building media literacy as public infrastructure so citizens can recognize and resist populist framing. None of these interventions suppresses speech, removes populist content from platforms, or prevents populist movements from existing. They achieve one thing only: they alter structural incentives so that populism no longer enjoys a systematic advantage through platform architecture.

The January 6 insurrection, with which this study opened, illustrates the stakes involved. The participants in that event were ordinary citizens who encountered a specific information environment on specific platforms, developed beliefs consistent with that environment, and acted on those beliefs—a trajectory that the structural conditions described in this study made more likely. Those structural conditions have not been resolved.

The path forward requires accepting three truths simultaneously: that technology shapes behavior in measurable ways that can be understood and altered; that genuine grievances underlie populist appeal and demand serious engagement; and that democracies possess the institutional capacity to reform media systems in ways that preserve freedom while reducing populism's structural advantage. This research provides the evidence and the analytical framework. What remains is the political will to act.

Limitations and Directions for Future Research:

Several important limitations of this study merit explicit acknowledgment, as they define both the boundaries of its conclusions and productive directions for subsequent inquiry. Most significantly, the Misinformation Susceptibility Survey (Experiment 2) was conducted exclusively among Korean high school students at a single institution. While this cohort is theoretically relevant—adolescents constitute a high-engagement, high-risk social media demographic—the sample's cultural and demographic homogeneity limits the generalizability of its findings. Susceptibility to populist framing is likely to vary meaningfully across national contexts, languages, educational systems, and political cultures. Future research should replicate this survey across diverse national samples—including populations in the United States, Western Europe, and other Asian democracies—to establish whether the observed patterns are robust across contexts or whether the Korean high school sample represents a specific case. Expanding the dataset in this manner would constitute both a methodological strengthening and a substantive contribution to cross-national comparative research on media literacy and populist susceptibility.

The Echo Chamber Simulation (Experiment 1) is similarly constrained by its single-session, single-platform design. The observation of algorithmic drift within 100 YouTube Shorts recommendations captures an initial polarization effect but does not model cumulative exposure over days, weeks, or months. Longitudinal studies tracking algorithmic recommendation patterns across extended periods and multiple platforms would provide a more complete account of how ideological drift develops in real-world conditions. Additionally, the Populist Rhetoric Content Coding Project (Experiment 3) focused primarily on Western democratic contexts; future work should extend this analysis to non-Western political environments, where the relationship between social media, populism, and democratic institutions may take structurally different forms. Addressing these limitations would significantly strengthen the evidentiary foundation for the policy recommendations advanced in this study and would deepen scholarly understanding of the mechanisms through which digital platforms interact with democratic processes globally.

First, that technology shapes behavior in measurable ways that can be understood and altered.

Second, that genuine grievances undergird populist appeal and that addressing these grievances is essential to democratic resilience.

Third, that democracies possess the capacity to reform media systems in ways that preserve freedom while reducing populism's structural advantage.

The alternative is to accept that democracy is incompatible with current social media architecture—that as platforms evolve toward greater algorithmic sophistication and reach, democracy degrades. If this is true, then democracies must choose: reform platforms or accept democratic decline.

We believe reform is possible. This research provides the evidence and the roadmap. What remains is the political will to act.

■ Acknowledgments

The authors would like to thank the teachers and academic mentors who provided methodological guidance and critical feedback during the development of this paper. We are particularly grateful for discussions that helped refine the experimental design and the framing of the policy analysis.

The authors would also like to thank the anonymous reviewers for their insightful comments and constructive feedback.

This research was conducted independently and did not receive any external funding. All errors and interpretations remain the responsibility of the authors.

■ References

1. Biswas, A., Javadian Sabet, A., & Lin, Y.-R. (2025). Toxic politics and TikTok engagement in the 2024 U.S. election. *Harvard Kennedy School Misinformation Review*, 6(4). <https://misinforeview.hks.harvard.edu/article/toxic-politics-and-tiktok-engagement-in-the-2024-u-s-election/>
2. Pew Research Center. (2025). How Americans' trust in information from news organizations and social media sites has changed over time. <https://www.pewresearch.org/short-reads/2025/10/29/how-americans-trust-in-information-from-news-organizations-and-social-media-sites-has-changed-over-time/>
3. Pew Research Center. (2025). Many Americans who generally distrust national news organizations still express trust in certain outlets. <https://www.pewresearch.org/short-reads/2025/09/02/many-americans-who-generally-distrust-national-news-organizations-still-express-trust-in-certain-outlets/>
4. Bialik, C., & Orth, T. (2025). Trust in media 2025: Which news sources Americans use and trust. *YouGov*. <https://today.yougov.com/politics/articles/52272-trust-in-media-2025-which-news-sources-americans-use-and-trust>
5. Ad Fontes Media. (2024). *Media bias chart* (Version 12.0). <https://adfontesmedia.com/media-bias-chart-jan-2024/>
6. Ad Fontes Media. (2024). *The media bias chart features 170 news and news-like sources*. <https://adfontesmedia.com/flagship-media-bias-chart-aug24/>
7. González-Aguilar, J. M., Segado-Boj, F., & Makhortykh, M. (2023). Populist right parties on TikTok: Spectacularization, personalization, and hate speech. *Media and Communication*, 11(2). <https://www.cogitatiopress.com/mediaandcommunication/article/view/6358>
8. Albertazzi, D., & Bonansinga, D. (2023). Populism goes TikTok: The German experience. *Media research working paper*. https://www.researchgate.net/publication/391988114_Populism_goes_TikTok_the_German_experience
9. Domínguez-García, R., Baptista, J. P., Pérez-Curiel, C., & Fonseca, D. E. M. d. (2025). From Emotion to Virality: The Use of Social Media in the Populism of Chega and VOX. *Social Sciences*, 14(5), 255. <https://doi.org/10.3390/socsci14050255>
10. Bösch, F. (2023). Alternative TikTok tactics: How the German right-wing populist party AfD plays the platform. *Unpublished manuscript*. https://www.researchgate.net/publication/374799634_Alternative_TikTok_Tactics_How_the_German_Right-Wing_Populist_Party_AfD_Plays_the_Platform
11. Eder, M., & Haßler, J. (2025). Does the TikTok feed lean right? Exposure to political party content among non-partisan users during regional and federal elections in Germany. *Preprint*. https://www.researchgate.net/publication/397203143_Does_the_TikTok_feed_lean_right_Exposure_to_Political_Party_Content_among_non-partisan_users_during_regional_and_federal_elections_in_Germany
12. Fondapol. (2024). AfD: The German far-right at a dead end. <https://www.fondapol.org/en/study/afd-the-german-far-right-at-a-dead-end/>
13. Prior, H. Social media and the rise of radical right populism in Portugal: the communicative strategies of André Ventura on X in the 2022 elections. *Humanit Soc Sci Commun* 11, 761 (2024). <https://doi.org/10.1057/s41599-024-03224-w>
14. Santos, H., & Roque, V. (2024). Populist leaders as gatekeepers: André Ventura uses news to legitimize the discourse. *Social Sciences & Missions*, 5(3), 84. <https://www.mdpi.com/2673-5172/5/3/84>
15. Dias, B. (2022). Political messianism in Portugal: The case of André Ventura. *Slovak Journal of Political Sciences*, 22(1). https://www.researchgate.net/publication/362464820_Political_Messianism_in_Portugal_the_Case_of_Andre_Ventura
16. Jaakkola, E., & Sakki, I. (2024). Influencer or radical right politician? Finnish populist radical right communication on TikTok. *University of Aberdeen, School of Social Science Blog*. <https://www.abdn.ac.uk/socsci/blog/influencer-or-radical-right-politician-finnish-populist-radical-right-communication-on-tiktok-blurs-the-line-between-the-two/>
17. Pettersson, K., & Sakki, I. (2024). Beyond negativity? A laboratory study on emotional responses to populist strategies on social media. *Frontiers in Political Science*. <https://www.frontiersin.org/articles/10.3389/fpos.2024.1415403>
18. Pettersson, K., & Sakki, I. (2025). Unveiling populist tactics on TikTok: A multimodal critical discursive psychology approach. In *Imagery of hate online* (Chap. 7). Open Book Publishers. <https://doi.org/10.11647/obp.0447.ch7>
19. CheckFirst, & Faktabaari. (2024). Finnish far-right videos highly recommended by YouTube during the presidential race. <https://checkfirst.network/checkfirst-and-faktabaari-present-new-study-showing-finnish-far-right-videos-highly-recommended-by-youtube-during-the-presidential-race/>
20. Rosa-Luxemburg-Stiftung Brussels. (2024). Northern exposure: Finland's far right between scandal and mainstream. <https://rosalux.eu/en/2024/die-extreme-rechte-in-finnland-zwischen-skandaloes-und-salonfaehig/>
21. Wahl-Jørgensen, K. (2024). Farage on TikTok: The perfect populist platform. *UK Election Analysis 2024*. <https://www.electionanalysis.uk/uk-election-analysis-2024/section-6-the-digital-campaign/farage-on-tiktok-the-perfect-populist-platform/>
22. The Citizens. (2024). The TikTok election: How Reform got to be so popular. <https://the-citizens.com/2024/07/the-tiktok-election-how-reform-got-to-be-so-popular/>
23. Freedom House. (2020). France: Freedom on the Net 2020 country report. <https://freedomhouse.org/country/france/freedom-net/2020>

■ Authors

Sejin Oh is a senior at Hankuk Academy of Foreign Studies in South Korea. She is interested in international relations, with a particular focus on education policy for adolescents and special education. She presents this paper in light of its growing relevance within an increasingly polarized global political climate.

Daeho Lee is a high school student who became interested in the political implications of digital platforms after observing the events of January 6th unfold online. He argues that algorithms are not neutral, and his research examines how platform design can facilitate the spread of populism. Through collaborative work, he explores structural reforms such as transparency, friction, and digital literacy as mechanisms to safeguard democratic systems.

Chehwan Kim is a senior at Hankuk Academy of Foreign Studies with academic interests in political science and economics. As a member of WSDC Korea (Top 9), he represents his country in international debate tournaments, bringing a strong background in argumentation and global affairs to his academic work.