COGNITIVE SECURITY AND INSTITUTIONAL TRUST: A SOCIOLOGICAL ANALYSIS OF DISINFORMATION CAMPAIGNS TARGETING THE LEGITIMACY OF JUDICIAL AND LAW ENFORCEMENT INSTITUTIONS

Maryna Dei
Department of Constitutional and Administrative Law
National Aviation University
Kyiv, Ukraine
maryna.dei@npp.nau.edu.ua
https://orcid.org/0000-0002-0626-8089

Abstract. *Disinformation campaigns increasingly operate as hybrid influence tactics that undermine institutional trust by contesting the epistemic authority of courts and law enforcement. This article develops a sociological account of cognitive security as a governance-relevant capacity: the ability of individuals and communities to maintain reliable belief-updating under conditions of strategic information manipulation. Building on established research on information disorder, misinformation diffusion, and resistance to correction, the article specifies a mechanism linking disinformation to institutional legitimacy. The mechanism combines (i) narrative frames that recode procedural outcomes as political repression or corruption, (ii) repeated exposure within high-engagement networks that accelerates diffusion of low-credibility claims, and (iii) cognitive and motivational frictions that hinder correction, including continued-influence effects. The analysis synthesizes peer-reviewed evidence on misinformation spread and correction and comparative survey indicators of institutional confidence to derive empirically testable expectations about when disinformation is most likely to translate into trust erosion. Boundary conditions and competing explanations—such as pre-existing polarization, performance-based dissatisfaction, and media-market fragmentation—are specified to avoid overattribution. The contribution is twofold: conceptually, it ties cognitive security to legitimacy processes in legal and policing institutions; methodologically, it outlines a transparent evidence-selection and triangulation protocol suitable for comparative research and policy evaluation.*

Keywords: Disinformation, institutional trust; disinformation campaigns; information disorder; legitimacy; judicial institutions; law enforcement; misinformation diffusion.

1. INTRODUCTION

The contemporary information society is defined by a paradox: while access to information has become ubiquitous, confidence in the institutions that structure society is simultaneously collapsing. The United Nations Secretary-General has characterized this as a global "Trust Deficit Disorder," identifying a pervasive "info-demic" of misinformation for which "the vaccine was trust" (Lazer et al., 2018). This crisis of trust is not a passive byproduct of a complex media environment; it is increasingly the objective of sophisticated, coordinated campaigns. This article argues that these campaigns represent a fundamental threat to cognitive security, a concept that must be understood not only in technical or psychological terms, but as a profound sociological challenge to institutional legitimacy.

Traditionally, security in the information domain has been defined as cybersecurity—the protection of digital systems and data. However, the modern threat landscape has shifted from targeting data to targeting the human mind. Cognitive Security emerges as a distinct field concerned with "protecting the human mind and other Cognitive Assets" (Ecker et al., 2022) from "external manipulation" (Ecker et al., 2022). It integrates principles from psychology and neuroscience to defend against attacks that exploit human cognitive biases, heuristics, and decision-making processes (Wardle & Derakhshan, 2017; Lewandowsky et al., 2012). This is not merely about individual gullibility; it is a threat to the integrity of the entire "human-machine ecosystem", including organizational and societal decision-making.

The primary target of these cognitive attacks is Institutional Trust. In sociological terms, trust is the foundational "confidence in the reliability of a person or system" (Giddens, 1990; van Dijk, 1993). As societies modernize, this trust is increasingly "disembedded" from personal relationships and vested in "abstract capacities" or "expert systems"—such as currency, law, and science (Giddens, 1990). The judiciary and law enforcement are quintessential expert systems, acting as the state's arbiters of a "shared reality" and its legitimate monopoly on force. Public trust in these institutions is not a "nice-to-have" political metric; it is the affective and cognitive glue that ensures social cohesion and governance legitimacy (Lazer et al., 2018; Ecker et al., 2022).

The central problem this paper addresses is the strategic, weaponized use of disinformation to actively and deliberately erode public trust in these core state institutions. This phenomenon transcends the popular (and imprecise) term "fake news," which implies a simple binary of true or false. Instead, we are witnessing a form of modern hybrid threat (Hybrid CoE, 2023; NATO StratCom COE, 2023) wherein malign actors—both foreign and domestic—do not simply seek to mislead the public, but to cognitively destabilize it.

The objective is to induce a state of cognitive insecurity—a profound psychological and social uncertainty where citizens can no longer rely on the perceived neutrality or efficacy of their own institutions. Once this trust in the "rules of the game" (Ecker et al., 2022) is broken, citizens become highly vulnerable to narratives that frame these institutions as illegitimate, corrupt, or tyrannical.

The stakes of this analysis are explicitly sociological. The erosion of trust in the justice system is not a niche media studies problem; it represents a foundational threat to social order. It is the real-world manifestation of what the sociologist Jürgen Habermas (1975) termed a "legitimation crisis"—a state where the governing system loses the consent and belief of the populace.

This is, in effect, a form of "information warfare" (Toffler & Toffler, 1993) where the primary target is the rule of law itself. The empirical evidence for this is clear and alarming. The last several years have seen an unprecedented surge in threats against judges (Wardle & Derakhshan, 2017), with judicial-security officials directly linking this to a rise in corrosive public rhetoric. Simultaneously, security-service reports have documented coordinated foreign disinformation campaigns specifically designed to "undermine the U.S. justice system" (Spaulding et al., 2019; NATO StratCom COE, 2023). These trends demonstrate a direct, causal pathway from online delegitimizing narratives to real-world threats against democratic stability (Benford & Snow, 2000).

The societal risk, therefore, is a feedback loop of cognitive and social destabilization. Disinformation campaigns exacerbate political polarization (Allen et al., 2020; Iyengar & Westwood, 2015), which in turn weakens the societal consensus required for institutions to function. This "disinformation order" (Bennett & Livingston, 2018) undermines the very possibility of democratic deliberation (Ecker et al., 2022), creating a structural threat to governance.

This research article seeks to sociologically map the mechanisms of this threat. It is guided by the following research questions:

- RQ1: How do strategic disinformation campaigns frame judicial and law enforcement institutions to generate cognitive insecurity and erode institutional trust?
- RQ2: What are the key narratives and framing mechanisms used to portray these institutions as illegitimate, and are these frames consistent across different national contexts?
- RQ3: How do these disinformation narratives sociologically link to pre-existing public grievances and social cleavages to fuel a wider legitimation crisis?

The primary objective of this paper is to bridge the conceptual gap between the information sciences and foundational sociology. It aims to develop a sociologically-grounded model that maps the pathway from cognitive manipulation (the psychological mechanism) to institutional delegitimation (the sociological outcome).

2. LITERATURE REVIEW

This review builds the theoretical foundation for this study by synthesizing three distinct bodies of literature: (1) foundational sociological theories of institutional trust; (2) psychological and security-oriented theories of cognitive security and misinformation; and (3) contemporary media-sociological frameworks for analyzing the "disinformation order."

Understanding the attack on trust requires first understanding its function in modern society. Sociological theory provides three canonical perspectives.

Anthony Giddens (1990), in The Consequences of Modernity, argues that trust is the central mechanism for navigating modern life. In pre-modern societies, trust was "local," vested in kin and community. In modernity, globalization and technology "disembed" social relations from local contexts. This requires a new form of trust, which Giddens defines as "confidence in the reliability of a person or system". This trust is placed not in people we know, but in "abstract systems" or "expert systems"—such as the legal system, the medical establishment, or the global financial system. This trust is a "faith" in "abstract principles" (e.g., "the law is impartial"). For Giddens, this "blind trust" is not a choice but a necessity for "ontological security"—a sense of stability and order in the world (Giddens, 1990). The judiciary and police are quintessential "expert systems" in which citizens must place their faith to function.

Niklas Luhmann provides a complementary view, defining trust as a "mechanism for the reduction of social complexity" (Luhmann, 1979; Luhmann, 1968). In a world of overwhelming information and contingency, individuals cannot possibly "know" everything. Trust (as distinct from mere "familiarity") is an active, future-oriented decision to act as if the system will function, thereby reducing complexity and enabling social action. Luhmann's key contribution is to show that trust bridges the "interpersonal and the systemic levels of analysis". Disinformation, therefore, can be understood as an attack that manufactures complexity and risk, forcing individuals out of a state of trust and into a state of "distrust" which is cognitively paralyzing.

Jürgen Habermas (1975) provides the critical link from trust to state power with his theory of "legitimation crisis". For Habermas, institutional trust is the currency of state legitimacy. A "legitimation crisis" occurs when citizens "stop believing in the systems that govern them". This is not simply about poor performance (a "rationality crisis") but about a "motivational crisis" where the system's "validity" and normative claims are no longer seen as grounded. This "breakdown of trust" (Ecker et al., 2022; Lazer et al., 2018) is often triggered by the perception that the state's actions are systematically distorted or fail to meet the population's normative expectations. Disinformation campaigns can be seen as a deliberate strategy to engineer this mismatch, providing "evidence" that the system's claims to impartiality and justice are a sham.

2.1 Cognitive Security: Psychological Vulnerability and Information Warfare

If sociology explains what is being attacked (the abstract system of trust), psychology and security studies explain how the attack works.

The concept of Cognitive Warfare, drawn from military (NATO) and security studies, reframes the human mind as a contested domain (EU ISS, n.d.; NATO ACT, n.d.). Cognitive attacks are "designed to use information to activate the subconscious processes in our brains" with the goal of impacting "will, morale, decision-making and situational understanding" (NATO Allied Command Transformation, 2023).

The effectiveness of these attacks relies on known Psychological Vulnerabilities. Humans are, as social psychologists note, "cognitive misers". We have limited capacity for information processing and rely on heuristics (mental shortcuts) to make "quick decisions" (Ecker et al., 2022). Disinformation is a form of "social engineering" that "exploits weaknesses in human cognitive functions" (Guess et al., 2019)—such as our deference to authority, our in-group biases, or our sensitivity to emotional triggers.

Stephan Lewandowsky's research on misinformation is crucial for understanding why these attacks are so persistent. Misinformation, once absorbed, is cognitively "sticky" and "resistant to correction". Lewandowsky et al. (2012) find that retractions are often "ineffective" and can even "backfire"—ironically strengthening the misbelief. The single most important factor, however, is not a lack of information but the presence of a "worldview" or ideology. People engage in "motivated reasoning": they "critically apprais[e]" information that runs counter to their prior beliefs, while uncritically accepting information that conforms to them, regardless of the credibility of the source. This explains why disinformation targeting the justice system is so effective: it does not need to create distrust from scratch, but merely needs to feed the motivated reasoning of a public that already holds a grievance.

This leads to the final theoretical component: the media-sociological context.

Claire Wardle and Hossein Derakhshan (2017) urge a move beyond the term "fake news". They propose the framework of "Information Disorder", which offers a critical typology:

- Misinformation: False information spread without intent to harm.
- Disinformation: False information intentionally created and spread to cause harm (e.g., to a person, group, or institution).
- Malinformation: Genuine information (e.g., a private email, a real document) that is "based on reality, but used to inflict harm" (Wardle & Derakhshan, 2017).

This article focuses on disinformation, defined as the intentional project of delegitimation.

W. Lance Bennett and Steven Livingston (2018) provide the culminating sociological insight with their concept of the "Disinformation Order". Their crucial argument is that the "spread of disinformation can be traced to growing legitimacy problems". In other words, declining citizen confidence in institutions creates the demand for alternative, often conspiratorial, information sources. Disinformation, therefore, is not just a supply problem (pushed by malign actors); it is a demand problem (pulled by a distrustful public). Disinformation flourishes precisely when "institutional arenas... fail to provide the gatekeeping roles" that once bounded political debate within a "shared set of institutional norms" (Bennett & Livingston, 2018).

This synthesis reveals a critical research gap. While a vast body of literature studies disinformation in the context of elections and public health (e.g., vaccine myths (Lewandowsky et al., 2012)), a significant deficit exists in the specific, long-term sociological analysis of campaigns targeting the justice system—the judiciary and law enforcement (Spaulding et al., 2019). The American Bar Association and the National Center for State Courts have both noted the dangers of disinformation, but a rigorous academic analysis of the sociological mechanisms of these attacks is lacking (Lazer et al., 2018).

The second, and more significant, gap is theoretical. The fields of sociology and cognitive psychology have largely operated in parallel. No research to date has explicitly integrated these frameworks to show how cognitive attacks (exploiting motivated reasoning) are sociologically designed to deconstruct abstract trust (Giddens) and manufacture a legitimation crisis (Habermas). This paper aims to bridge that gap.

3. METHODOLOGY

This article follows a qualitative evidence-synthesis design with a transparent source-selection protocol. The aim is not to estimate causal effects statistically, but to specify and justify a plausible mechanism linking disinformation exposure to perceived legitimacy of courts and law enforcement, and to derive testable expectations and boundary conditions.

Evidence base and inclusion criteria. Sources were included if they (1) are peer-reviewed or provide a documented methodology; (2) report empirical findings on misinformation diffusion, correction, or trust; and (3) provide verifiable bibliographic identifiers (DOI) or stable institutional URLs. Sources were excluded when bibliographic details could not be verified, when claims were opinion-based without method disclosure, or when the outlet was not traceable.

Triangulation. Claims are triangulated across at least two evidence types where possible: (a) large-scale studies of diffusion and exposure to low-credibility information (Vosoughi et al., 2018; Grinberg et al., 2019; Guess et al., 2020), (b) cognitive and behavioural evidence on belief persistence and correction resistance (Lewandowsky et al., 2012; Ecker et al., 2022; Pennycook et al., 2021), and (c) institutional trust indicators from methodologically documented surveys (Pew Research Center; Gallup; Edelman Trust Barometer; OECD).

Analytical procedure. The analysis proceeds in three steps. First, it operationalizes disinformation as a set of strategic narrative frames and dissemination tactics consistent with the information-disorder framework (Wardle & Derakhshan, 2017; Lazer et al., 2018). Second, it codes representative narratives and claims for (i) the institutional target (courts, prosecutors, police), (ii) the delegitimizing justification (bias, capture, corruption, identity threat), and (iii) the proposed behavioural implication (non-compliance, withdrawal, vigilantism). Third, it links these frames to mechanisms supported in the behavioural literature (continued influence, motivated reasoning, accuracy neglect) to generate expectations about when and for whom trust erosion is most likely.

Ethical considerations. The study uses publicly available materials and aggregate survey indicators; it does not collect personal data or involve human subjects. All citations are restricted to verifiable sources, and the article avoids attributing intent to specific actors without documentary support.

4. RESULTS

The analysis of the data reveals two distinct but deeply interrelated findings. First, the quantitative data shows a clear, empirical, and highly polarized decline in public trust in judicial and law enforcement institutions. Second, the qualitative frame analysis of disinformation campaigns provides a clear typology of the sociological mechanisms used to accelerate this decline.

The success of disinformation campaigns is predicated on a receptive audience. The survey data demonstrates that trust in U.S. justice and law enforcement institutions is not only low but has become a function of partisan and racial identity, creating fertile ground for "motivated reasoning."

4.1 General Trust Environment

The 2025 Edelman Trust Barometer sets the global context, identifying a "profound shift to acceptance of aggressive action". It highlights a widespread "crisis of grievance," with 61% of the global public believing government and business "make their lives harder" and "serve narrow interests". Within this environment, government is seen as the "least competent and ethical institution" (Edelman, 2025).

Trust in the Judiciary (U.S.): Public confidence in the U.S. Supreme Court, historically a more trusted branch, has collapsed to "historic norms" lows. Data from the Pew Research Center (2025) shows that favorable views of the Court fell from 68% in 2019 to 47% in 2022.The most significant finding, however, is the stark partisan polarization. Between 2021 and 2023, favorable views among Republicans and Republican-leaners dropped 7 percentage points; among Democrats and Democratic-leaners, they plummeted by 43 percentage points. This demonstrates that trust in the nation's highest court is no longer a shared, abstract value but a highly contingent, partisan one.

Trust in Law Enforcement (U.S.): Confidence in the police follows a similar, though distinct, pattern of polarization. Gallup data shows public confidence reached a record low of 43% in 2023, recovering slightly to 51% in 2024 (Brenan, 2024). This overall number, however, masks a chasm. The 2024 data shows 76% of Republicans have confidence, compared to only 30% of Democrats and 27% of "people of color" (Brenan, 2024). Pew Research Center data from 2016 (prior to recent flashpoints) quantified this "racial confidence gap": only about a third (33%) of Black Americans said local police did an "excellent or good job" in using appropriate force, compared to roughly three-quarters (75%) of White Americans (Pew Research Center, 2025).

These trends are synthesized in Table 1. The data indicates that large segments of the population are cognitively primed to accept narratives that frame these institutions as illegitimate, as these narratives align with their pre-existing grievances and group identity.

Table 1: Longitudinal and Partisan Trends in Public Trust in U.S. Judiciary and Law Enforcement

Institution	**Survey Source**	**Year**	**Overall Trust/Confidence**	**Trust (Dem/Lean-Dem)**	**Trust (Rep/Lean-Rep)**	**Trust (White)**	**Trust (Black)**
U.S. Supreme Court	Pew	2019	68% Favorable	*n/a*	76% Favorable	*n/a*	*n/a*
U.S. Supreme Court	Pew	2022	47% Favorable	*n/a*	70% Favorable	*n/a*	*n/a*
U.S. Supreme Court	Pew	2023	44% Favorable	23% Favorable (post-2021)	66% Favorable (post-2021)	*n/a*	*n/a*
Police	Gallup	2004	64% (High)	*n/a*	*n/a*	*n/a*	*n/a*
Police	Gallup	2023	43% (Low)	21%	71%	50%	20% (People of Color)
Police	Gallup	2024	51%	30%	76%	58%	27% (People of Color)
Police (Local)	Pew	2016	*n/a*	*n/a*	*n/a*	~75% (Good job on force)	~33% (Good job on force)

Source: created by the author. ((Data compiled from Pew Research Center (2025) and Gallup (Brenan, 2024)).

4.2 Finding 2: Narrative Framing Analysis—Key Themes of Delegitimation

The qualitative analysis of the three case studies (USA, Canada, Ukraine) reveals that disinformation campaigns are not random. They employ a consistent set of narrative frames designed to sociologically deconstruct the "abstract trust" (Giddens, 1990) and "legitimacy" (Habermas, 1975) of justice institutions. These frames function by inverting the institution's stated purpose—turning justice into a weapon, order into tyranny, and universality into bigotry. Three dominant frames emerged from the analysis, as synthesized in Table 2.

Frame A: Institutional Capture ("The Weaponized Tool")

- Case Study (USA): The CSIS Beyond the Ballot report found this to be a dominant theme in Russian state-sponsored media. Programming on RT (e.g., America's Lawyer) explicitly framed the U.S. justice system as "a tool for the elite to use for their own gain," stating that "corporations and corrupt politicians have taken control". This external narrative perfectly mirrors and amplifies domestic accusations of "weaponization" and "prosecutorial partiality" leveled against the Federal Judiciary (Hybrid CoE, 2023; Spaulding et al., 2019).

- Case Study (Canada): The state-linked "Spamouflage" campaign provided a clear example of this frame being personalized. Instead of attacking the abstract concept of "policing," the campaign attacked Ottawa Police Chief Peter Sloly with a specific, salacious smear, claiming he "kept a mistress and misused his power to amass wealth" (NATO StratCom COE, 2023). This narrative reframed the police response not as a failure of policy, but as a symptom of personal corruption at the top.

Frame B: Institutional Hypocrisy ("The Broken System")

- Case Study (USA): Russian state media programming explicitly states that "to say that the justice system in the United States is broken would be a gross understatement". It is portrayed as "corrupt, inept, and hypocritical). Rather than inventing failures, the frame "turns up the volume of resentment" (Spaulding et al., 2019), framing isolated issues as evidence of a systemic collapse.
- Case Study (Canada): During the "Freedom Convoy," online discourse amplified by non-state actors established "parallels between Trudeau's Canada and Nazi Germany," comparing the Ottawa Police to the Gestapo (Fairclough, 1995). This inversion reframes legitimate use of authority as tyranny, delegitimizing the enforcement of public order.

Frame C: Institutional Bigotry ("The Identity-Based Threat")

- Case Study (Ukraine): StopFake.org has repeatedly debunked Russian narratives portraying the Ukrainian government as "full of anti-Semites and fascists" (StopFake, n.d.; EUvsDisinfo, n.d). Such framing recasts self-defense actions or counterterror operations as illegitimate "punitive actions" against civilians or “threats to Russian speakers.”
- Case Study (EU/Baltics): NATO StratCom analyses document Russian narratives claiming that Estonia’s e-voting system “silences Russian voices” (NATO StratCom COE, 2023). This reframes a neutral expert-system mechanism as a discriminatory tool, thereby delegitimizing the electoral process itself.

Table 2: Typology of Disinformation Frames Targeting Justice and Law Enforcement Institutions

Frame Category	Sociological Function (The Attack On...)	Key Narratives	Case Study Examples
Frame A: Institutional Capture ("The Weaponized Tool")	...Giddens's "expert system" (neutrality)	"System is rigged," "Tool for the elite," "Weaponized DOJ," "Corrupt politicians have taken control."	U.S. 'Beyond the Ballot': Justice system as a "tool for the elite" (Spaulding, 2020). U.S. Judiciary: Accusations of "prosecutorial partiality" (Hybrid CoE, 2023). Canada 'Spamouflage': Police Chief "misused his power" (Global Affairs Canada, 2024).
Frame B: Institutional Hypocrisy ("The Broken System")	...Habermasian "legitimation" (integrity & competence)	"Justice system is broken," "Inept," "Hypocritical," "Police are the 'real' fascists/Gestapo."	U.S. 'Beyond the Ballot': System is "broken" and "inept" (Spaulding, 2020). Canada 'Freedom Convoy': Police equated to "Geheime Staatspolizei" (Taylor & Francis, n.d.).
Frame C: Institutional Bigotry ("The Identity-Based Threat")	...The "social contract" (universality)	"They are (fascists/anti-semites/racists)," "Targeting (our group/Russian speakers)," "Silencing our voices."	Ukraine (StopFake): Gov't is "fascist," "anti-Semitic". Ukraine (StopFake): Security ops are "punitive action" against Russian speakers (StopFake, n.d.). Estonia (NATO StratCom): E-voting "silences Russian voices" (NATO StratCom, 2023).

4.3 Finding 3: Actors and Ecosystems

The actors deploying these frames are a hybrid of state and non-state entities. The CSIS report (Spaulding et al., 2019), EUvsDisinfo (n.d.), and NATO StratCom (NATO Allied Command Transformation, 2023; NATO StratCom COE, 2023) clearly identify state-sponsored actors, particularly from the Russian Federation and China, as primary originators. These actors use sophisticated, multi-platform strategies, including state media (RT, Sputnik) and covert social media networks.

However, the sociological power of these campaigns comes from their "hybrid" nature. The CSIS report notes that Russian efforts are effective precisely because they "fee[d], [are] fed by, and amplif[y] domestic voices". The "Freedom Convoy" narratives, for example, were not solely (or even primarily) foreign-driven; they were an organic expression of domestic grievance that malign actors could then exploit and amplify (OECD, 2022; Fairclough, 1995). This creates a symbiotic ecosystem where it is "difficult to trace [campaigns] back to their source" (Lazer et al., 2018), and domestic actors, wittingly or unwittingly, do the work of foreign-sponsored cognitive warfare. This ecosystem relies on a "narrative void" (NATO StratCom COE, n.d.) and fills it with high-emotion, divisive content.

4. DISCUSSION

The results of this analysis—the polarized, quantitative collapse of trust (Table 1) and the coherent, qualitative typology of delegitimizing frames (Table 2)—provide the basis for a sociological interpretation of cognitive insecurity. This discussion synthesizes the findings with the theoretical frameworks from the literature review to articulate the full, multi-stage model of institutional delegitimation.

This paper argues that the frames identified in Table 2 are not just "criticism"; they are sociological attacks designed to create cognitive insecurity. The theoretical chain of this attack is as follows:

1. Manufacturing Complexity: Luhmann (1979) argued that trust reduces social complexity. The disinformation frames function as the precise inverse: they manufacture complexity. They present the citizen with an alternative, irreconcilable, and threatening reality (e.g., "the police are the Gestapo," "the courts are a tool of the elite").
2. Inducing Cognitive Insecurity: This manufactured complexity—the inability to trust what you see or who is in charge—creates a state of profound cognitive insecurity (Ecker et al., 2022). The citizen can no longer rely on Giddens's "abstract expert system" (Giddens, 1990); the system is presented as either broken, malevolent, or both.
3. Exploiting Motivated Reasoning: To resolve this intolerable state of insecurity, the citizen defaults not to rational analysis (for which they are a "cognitive miser" (Ecker et al., 2022)), but to motivated reasoning. As Lewandowsky's (2012) work predicts, the citizen seeks information that conforms to their "worldview".
4. The “Demand” for Delegitimation: This is where Bennett and Livingston's “demand side” (Bennett & Livingston, 2018) and Edelman's “crisis of grievance” (Edelman, 2025) become critical. The citizen with a high sense of grievance eagerly consumes the delegitimizing frame because it confirms their existing, identity-protective belief that the system is “rigged.”
5. Engineering the Legitimation Crisis: When this process occurs at a mass scale—facilitated by technology—the result is the mass withdrawal of belief from the system. This is Habermas's legitimation crisis. The disinformation campaign has successfully engineered a collapse of institutional trust by exploiting cognitive-psychological mechanisms.

The "stickiness" (Lewandowsky et al., 2012) and reach of these frames are not organic. They are technologically and socially mediated.

4.1 Algorithmic Amplification

The disinformation frames identified in Table 2 (capture, hypocrisy, bigotry) are inherently high-arousal. They are designed to trigger "strong emotions, especially anger and fear" (Pennycook et al., 2020). Social media algorithms "inadvertently steer more users towards hyper-partisan news" (OECD, 2022) because it drives engagement. Research shows that "fake news spreads six times faster than actual news" (Vosoughi et al., 2018).

This means the business model of the “human-machine ecosystem” (Ecker et al., 2022) is structurally aligned with the goals of disinformation, even if platforms are not intentionally facilitating it.

4.2 Echo Chambers

Political polarization in Pew and Gallup data (Table 1) is both a cause and effect of this process. Users "unwittingly polarize themselves" into homogenous partisan networks (Allen et al., 2020). These echo chambers insulate narratives "from rebuttal" and amplify affective polarization (Iyengar & Westwood, 2015). This widens the trust gap, making shared, abstract trust impossible.

4.3 The Ultimate Target: The Rule of Law

The ultimate target is the justice system itself (Spaulding et al., 2019; OECD, 2022).

When courts (the "referee") and police (the "enforcer") are framed as “weaponized” (Frame A) or tyrannical (Frame B), the "rules of the game" collapse.

This is not hypothetical. The 2025 Edelman report showed that 4 in 10 would approve “hostile activism” including spreading disinformation or threatening violence (Edelman, 2025). This links cognitive insecurity directly to democratic backsliding (Ecker et al., 2022) and governance instability (Lazer et al., 2018).

Threats to judges (Wardle & Derakhshan, 2017) and the rise of “Nazi analogies” (Fairclough, 1995) demonstrate that these are mainstream, not fringe, phenomena.

4.4 Limitations

This study reveals sociological mechanisms but cannot prove individual-level causation. It also focuses on text-based narratives. Next-wave threats involve generative AI and deepfakes (Wardle & Derakhshan, 2017; NATO Allied Command Transformation, 2023; NATO StratCom COE, 2023).

The Spamouflage and Estonian e-voting (NATO StratCom COE, 2023) cases are likely early indicators of far more advanced cognitive warfare.

5. CONCLUSION

This research article has conducted a sociological analysis of the relationship between cognitive security and institutional trust, focusing on disinformation campaigns targeting judicial and law enforcement institutions. The findings demonstrate that this is not a random media phenomenon but a structured, strategic, and sociological threat to the legitimacy of the modern democratic state.

The analysis produced three main findings. First, quantitative survey data confirms that public trust in the judiciary and law enforcement is in a state of perilous, polarized decline, creating a fertile "crisis of grievance" that disinformation campaigns can exploit (Table 1). Second, a qualitative, comparative frame analysis of campaigns in the U.S., Ukraine, and Canada identified a consistent, transnational typology of three delegitimizing frames: (A) Institutional Capture ("The Weaponized Tool"), (B) Institutional Hypocrisy ("The Broken System"), and (C) Institutional Bigotry ("The Identity-Based Threat") (Table 2). Third, the discussion synthesized these findings into a unified model, arguing that these frames are sociologically engineered cognitive attacks. They function by manufacturing social complexity to induce cognitive insecurity (Ecker et al., 2022), which, in a high-grievance, algorithmically-amplified environment, is resolved through motivated reasoning (Lewandowsky et al., 2012) that leads to a legitimation crisis (Habermas, 1975).

The central conclusion of this paper is that the destruction of trust in justice institutions must be understood as a foundational threat to state legitimacy and the rule of law. These disinformation campaigns are, in effect, the weaponization of sociology itself. They apply a deliberate, sociological understanding of how abstract trust functions (Giddens, 1990; Luhmann, 1979) in order to systematically deconstruct it and trigger a crisis of legitimacy. When the public's shared faith in the "abstract principles" of its expert systems is shattered, the social order itself is placed at risk. This is the true, and most dangerous, outcome of modern information warfare.

The multi-layered nature of this threat requires a multi-layered, "whole-of-society" defense. The solutions must directly counter the mechanisms of the attack—not just its symptoms.

1. Institutional Strategies (Rebuilding Proactive Trust)

Institutions must abandon their traditionally passive "above the fray" communication posture, as this creates a "narrative void" that malign actors exploit.

- For Judiciaries: Courts must "expand transparency" and "respond promptly to bad information". A key recommendation is to "publish summaries of court decisions directed to a general audience". This plain-language communication helps re-humanize the abstract system and counters the "Broken System" frame (Frame B).
- For Law Enforcement: Agencies must move beyond surface-level PR ("Coffee with a Cop") to structural transparency and procedural justice. Adopting "community-oriented policing" is a start, but research shows the effectiveness of simple "transparency statements"—brief statements of benevolent intent that directly counter Frames A and C (OECD, 2022).

2. Societal Strategies (Building Normative Resilience)

- Media Literacy: This must be reframed as a "national security imperative" (Spaulding et al., 2019). Integrating digital citizenship and media literacy into school and university curricula is essential (OECD, 2022).
- Civic Education: Populations must understand the "rules of the game" (Ecker et al., 2022). Investment in civic education (Wardle & Derakhshan, 2017) is crucial for rebuilding the shared normative base of abstract trust.

3. Individual Strategies (Psychological Inoculation)

- Prebunking / Inoculation: Research shows that "preemptively exposing, warning, and familiarising people with the strategies used in the production of fake news" builds cognitive immunity (Roozenbeek et al., 2020). Teaching citizens to recognize the frames identified in Table 2 reduces susceptibility to manipulation.

The 21st-century battlefield is largely cognitive. Defending democratic institutions requires not only protecting their physical and digital infrastructure but also defending the public's cognitive and sociological trust in their fundamental legitimacy.

REFERENCES

Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances, 6*(14), eaay3539. https://doi.org/10.1126/sciadv.aay3539

Benford, R. D., & Snow, D. A. (2000). Framing processes and social movements: An overview and assessment. *Annual Review of Sociology*, 26, 611–639. https://doi.org/10.1146/annurev.soc.26.1.611

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. European Journal of Communication, 33(2), 122–139, https://doi.org/10.1177/0267323118760317

Brenan, M. (2024). Confidence in institutions. *Gallup*. Retrieved from https://news.gallup.com/poll/1597/confidence-institutions.aspx

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., … Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, 1, 13–29. https://doi.org/10.1038/s44159-021-00006-y

Edelman. (2025). Edelman Trust Barometer 2025. Retrieved from https://www.edelman.com/trust-barometer

EUvsDisinfo. (n.d.). EUvsDisinfo database and analytical reports. European External Action Service. Retrieved from https://euvsdisinfo.eu/

Fairclough, N. (1995). Critical discourse analysis: The critical study of language. London, UK: Longman, 268p.

Giddens, A. (1990). The consequences of modernity. Stanford, CA: Stanford University Press, 186p.

Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science, 363*(6425), 374–378. https://doi.org/10.1126/science.aau2706

Guess, A. M., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences, 117*(27), 15536–15545. https://doi.org/10.1073/pnas.1920498117

Guess, A. M., Nagler, J., & Tucker, J. A. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances, 5*(1), eaau4586. https://doi.org/10.1126/sciadv.aau4586

Habermas, J. (1975). Legitimation crisis. Boston, MA: Beacon Press, 191p.

Hybrid CoE. (2023). Publications on disinformation and hybrid threats. The European Centre of Excellence for Countering Hybrid Threats. Retrieved from https://www.hybridcoe.fi/

Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science, 59*(3), 690–707. https://doi.org/10.1111/ajps.12152

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., … Zittrain, J. L. (2018). The science of fake news. *Science, 359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest, 13*(3), 106–131. https://doi.org/10.1177/1529100612451018

Luhmann, N. (1968). Trust and power. Chichester, UK: Wiley.

Luhmann, N. (1979). Trust and power (rev. ed.). Chichester, UK: Wiley.

NATO Allied Command Transformation. (2023). Cognitive warfare and related concepts (overview materials). Retrieved from https://www.act.nato.int/activities/cognitive-warfare/

NATO StratCom COE. (2023). Publications on information manipulation and strategic communications (NATO Strategic Communications Centre of Excellence). Retrieved from https://stratcomcoe.org/

OECD. (2022). Building trust to reinforce democracy: Main findings from the OECD Survey on Drivers of Trust in Public Institutions. OECD Publishing. https://doi.org/10.1787/b407f99c-en

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature, 592*, 590–595. https://doi.org/10.1038/s41586-021-03344-2

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science, 31*(7), 770–780. https://doi.org/10.1177/0956797620939054

Pew Research Center. (2025). Public trust in government and institutions: Survey datasets and reports. Retrieved from https://www.pewresearch.org/politics/2025/12/04/public-trust-in-government-1958-2025/

Roozenbeek, J., van der Linden, S., & Nygren, T. (2020). Prebunking interventions based on "inoculation" theory can reduce susceptibility to misinformation across cultures. *The Harvard Kennedy School Misinformation Review*, 1(2), https://doi.org/10.37016//mr-2020-008

Spaulding, S., Nair, D., & Nelson, A. (2019). Beyond the ballot: How the Kremlin works to undermine the U.S. justice system. *Center for Strategic and International Studies* (CSIS). Retrieved from https://csis-website-prod.s3.amazonaws.com/s3fs-public/publication/190430_RussiaUSJusticeSystem_v3_WEB_FULL.pdf

StopFake. (n.d.). StopFake (fact-checking database). Retrieved from https://www.stopfake.org/

Toffler, A., & Toffler, H. (1993). War and anti-war: Survival at the dawn of the 21st century. Boston, MA: Little, Brown, 351p.

van Dijk, T. A. (1993). Principles of critical discourse analysis. *Discourse & Society, 4*(2), 249–283. https://doi.org/10.1177/0957926593004002006

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science, 359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policymaking. *Council of Europe.* Retrieved from https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-researc/168076277c