Smart Technology Investments

Cognitive Wars: the AI Industrialization of Influence

Nov 01–Nov 08, 2025 | Sources: 11 | Anchor Status: Anchor-Absent | Report Type: Theoretical Research | Horizon: Near-term | Confidence: 0.480 [*]

SD	AC	MT	RR
0.80	0.00	0.45	0.65

Alignment: 6.0   Theory Depth: 6.0   Clarity: 7.0

> **Disclosure & Method Note:** This is a *theory-first* brief. Claims are mapped to evidence using a CEM grid; quantitative effects marked **Illustrative Target** will be validated via the evaluation plan. Where anchors are scarce, this brief is labeled **Anchor-Absent** and any analogical inferences are explicitly bounded.

Abstract & Theory-First Framing.

Outline

Introduction: theory-first framing

A theory-first approach foregrounds causal mechanisms—how material and institutional changes causally enable new forms of cognitive combat—before seeking broad empirical generalization. This avoids descriptive conflation of distinct phenomena (propaganda, public diplomacy, psychological operations, computational propaganda) and provides analytical leverage to predict continuities and ruptures as technologies evolve.

Central research question: How did industrialization influence the emergence, scale, and techniques of cognitive wars, and what does that historical relationship imply for the ongoing AI-driven industrialization of influence?

Two methodological desiderata follow: (1) prioritize mechanism identification (transport, communications, production, bureaucratization) using process tracing; (2) triangulate with quantitative proxies (print runs, literacy, telegraph traffic) where available to bound causal claims.

Theoretical Framework: cognitive wars

Definition: "Cognitive wars" are strategic contests that target human perception, beliefs, decision nodes, and collective cognition with the explicit intention to change behavior or political outcomes. They differ from kinetic war in objectives and mediating means: outcomes are cognitive states and social coordination rather than territorial control, though both interact.

Interdisciplinary situating: Cognitive wars draw on political theory of war (strategy and objectives), cognitive science (belief formation, attention, memory), information theory (signal, noise, channel capacity), and sociology (collective behavior, institutions). Theoretical model links material infrastructures to three cognitive capacities:

- Reach: the set and scale of audiences addressable by messaging.
- Salience & Dose: intensity and frequency of exposures required to change cognitive states.
- Persistence & Fidelity: durability of changed beliefs and accuracy/consistency of messages across channels.

Conjecture: Industrialization increased reach, reduced marginal cost per exposure, standardized message fidelity, and institutionalized production—transformations that permitted sustained campaigns with measurable population-level cognitive effects.

Foundations: canonical anchors and selection strategy

Why these anchors?

- Selection strategy: anchors are chosen across abstraction layers to ensure both domain specificity and theoretical depth. Direct sources (Layer 1) address contemporary mechanisms and empirical patterns of computational propaganda and influence campaigns. Domain sources (Layer 2) provide strategic and historical context for information operations. Foundational sources (Layers 3–5) offer canonical models—network diffusion, information theory, learning theory, and social psychology—that ground causal inferences from material infrastructures to cognitive outcomes.

- Current anchor availability: the source corpus supplied for this brief is dominated by contemporary, accessible arXiv/preprint literature and domain reports (digital-era analyses). There are presently zero (0) peer-reviewed, non-preprint anchors among the supplied sources; this brief therefore records that gap and uses the available preprints as provisional empirical anchors while recommending targeted integration of canonical peer-reviewed sources (e.g., Granovetter on thresholds and weak ties; Shannon on channel capacity; Vapnik on learning theory) in subsequent journal submissions and policy dossiers.

- Practical rationale: when direct peer-reviewed anchors are missing, canonical papers from broader layers provide first-principles machinery to translate modern digital findings into theory; they supply mathematical and conceptual scaffolding for generalization beyond case-bound observations.

Selected (available) anchors used below include contemporary computational-propaganda and diffusion studies to ground claims about automated scaling and message dynamics [6][8][2]. Network-theoretic and social-diffusion preprints inform the mapping from infrastructure to cognitive reach [1][9][10]. Studies of rumor, fake news, and community dynamics support claims about persistence and echoing [7][11].

Theoretical Grounding and Conceptual Framework

Abstraction layers and concepts

- Layer 1 (Specific): automated influence operations—botnets, coordinated accounts, generative media—operate as instruments of cognitive targeting capable of producing high dose and personalization at low marginal cost [6][3][4].

- Layer 2 (Domain): information warfare and psychological operations frame objectives, legal/ethical constraints, and the military-political logic of influence campaigns [2].

- Layer 3 (Cross-disciplinary): diffusion and social influence models (thresholds, cascades, weak ties) explain how messages propagate unevenly across networks and how local exposure thresholds produce macro-level cascades [9][10].

- Layer 4 (Theoretical): information theory (channel capacity, noise), game theory (adversarial signaling), and systems theory (feedback and amplification) provide mechanisms by which infrastructures change the effective signal-to-noise ratio and enable closed-loop optimization.

- Layer 5 (Foundational): statistical learning and algorithmic bias explain how automated influence systems generalize from training data, encode biases, and adapt through optimization—critical for understanding AI-driven personalization and adversarial exploitation.

Reasoning chain from foundational principles to the specific topic

1. Information theory: industrial communication channels (print runs, telegraph lines, broadcast) increase effective channel capacity and reduce costs per bit, enabling greater throughput of persuasive content.
2. Network models: increased connectivity and centralized platforms create low-resistance paths for diffusion; threshold dynamics convert local persuasion into large-scale cascades [9][10].
3. Social psychology: exposure frequency, messenger credibility, and social reinforcement shape belief revision and persistence—conditions made tractable by standardized messaging and repeatability.
4. Statistical learning: automation and algorithmic targeting convert aggregated behavioral data into micro-targeted messaging primitives; iterative optimization closes the loop between outputs and influence effectiveness [3][4].

How tangential canonical papers contribute

Canonical works from broader layers (e.g., threshold models, weak ties, Shannon) are not specific to propaganda but define how information propagates and how channels impose limits. These formal results ground claims that material changes (e.g., capacity, latency) causally alter diffusion, persuasion thresholds, and systemic vulnerability—providing a rigorous bridge from infrastructure to cognition.

Conceptual map: material infrastructures → channel properties (capacity, latency, fidelity) → network diffusion dynamics (reach, speed, cascade thresholds) → cognitive outcomes (belief change, decision influence, identity shifts).

Literature Review: wars and industrialization

Existing scholarship on industrialized war emphasizes logistics, mass mobilization, and technology (rail, telegraph, mass conscription) but often treats informational effects as consequential but secondary. Computational-propaganda research documents contemporary tactics and automation [6][2], and the fake-news literature models diffusion differentials between true and false information [8][7]. However, a gap persists: studies seldom theorize industrialization as a structural cause of cognitive modalities of war across historical epochs. Information-warfare treatments frequently analyze techniques and platforms without tracing long-term structural conditions—mass education, standardized production, and bureaucratic routines—which enabled durable cognitive campaigns.

This brief connects historical and contemporary literatures by translating diffusion and systems-theoretic results into a causal narrative: industrial infrastructures lowered marginal cost and increased coordination capacity, which in turn changed the scale and persistence of influence operations.

Historical Context: industrialization and military transformation

Claim: Industrialization reconfigured communications, transport, and media ecosystems in ways that altered the feasible scale and duration of cognitive operations.

Mechanisms in the 19th and early 20th centuries: telegraphy reduced latency of cross-regional coordination; railways enabled synchronized mobilization and distribution of printed material; inexpensive printing and later mass broadcast standardized messaging; mass schooling created literate publics reachable at scale. Bureaucratization produced professional propaganda offices, routinizing message production and targeting.

Patterns: (1) infrastructural enabling—faster, wider dissemination; (2) institutionalization—permanent propaganda arms within states and parties; (3) measurement primitives—circulation figures, literacy rates provided crude metrics to calibrate campaigns. These patterns enabled sustained homefront mobilization and colonial governance through cognitive means as much as they enabled kinetic mobilization.

Conceptual Section: defining cognitive warfare

Differentiation: Cognitive warfare differs from conventional kinetic warfare in its primary objects (cognition, decision-making) and from narrower propaganda in its scale, persistence, and integration with material infrastructures and organizational routines.

Targets: individual beliefs and attitudes, group identity and cohesion, decision nodes (e.g., electoral publics, military command chains), and information environments (norms, trust metrics).

Temporal qualities: durable influence (institutional embedding, norm change) versus transient influence (short-term persuasion, tactical misinformation). Metrics for analysis:

- Reach: proportion of target population exposed (absolute and weighted by influence centrality).
- Dose: message frequency per unit time; cumulative exposures required for threshold crossing.
- Fidelity: consistency and accuracy of message content across channels.
- Institutional embedding: degree to which messaging is routinized within organizations (professional units, budgets, legal authorities).
- Feedback loops/amplification: presence and strength of automated monitoring/optimization that closes the influence loop.

Mechanisms: industrialization's influence on cognitive operations

Unique claim set (not repetition of summary): Industrial-era technologies changed five mechanism classes relevant to cognitive warfare:

1. Channel capacity and latency: Telegraph and postal rail networks increased information throughput and decreased delivery time, permitting synchronous narrative control over wide areas and enabling coordinated cognitive operations across fronts [10].

1. Standardization and repeatability: Printing presses and later broadcast permitted identical message replication at scale. This lowered variance in message fidelity, making dose-response calibration tractable.

1. Audience legibility: Mass schooling and census/bureaucratic records produced legible target populations (demographics, literacy), enabling segmentation and targeted appeals.

1. Professionalization and economies of scale: Dedicated propaganda bureaus, party presses, and media firms institutionalized message production, enabling prolonged campaigns and iterative improvement (A/B-like testing by circulation and feedback about reactions).

1. Closed-loop coordination: Bureaucratic reporting and media monitoring created feedback channels—early measurement primitives that allowed campaigns to adapt in response to signal effectiveness and social reaction patterns. Modern AI closes these loops through automated analytics and personalization [3][4].

Each mechanism materially altered the cost function of influence: reduced marginal cost per exposure, increased effective exposures per unit time, and increased marginal return on coordination investments.

Case Studies: industrial-era cognitive campaigns

Comparative historical examples reveal recurring patterns of infrastructural enabling, institutionalization, and measurable cognitive shifts.

1. 19th-century nationalism and print culture: mass-circulation newspapers and serialized novels created shared imaginaries and reinforced national identity across disparate regions; mass literacy increased receptivity to standardized narratives.

1. Colonial counterinsurgency propaganda: bureaucratic pamphlet campaigns and controlled press in imperial territories used standardized messaging paired with administrative records to target and reshape local political identities, though success varied with cultural legibility and local networks.

1. World War I homefront mobilization: state-run ministries of information, censorship regimes, and mass posters coordinated to sustain morale and redirect dissent. These campaigns relied on industrial production (print runs, film) and centralized distribution.

Limits and variation: industrial capacity did not guarantee cognitive success. Failures occurred where messaging mismatched local cultural schemas, where rival infrastructures (underground presses, oral networks) resisted standardization, or where feedback loops were absent (no metricization to adjust campaigns). These limits underscore the interaction between infrastructure and social context.

Methodology

Design: mixed-methods, comparative-historical approach.

- Qualitative: archival research (propaganda office records, circulation reports), discourse analysis, and process tracing to identify mechanisms and sequencing.
- Quantitative: proxies to operationalize industrial measures—print runs, telegraph/parcel traffic volumes, literacy rates, and, where available, circulation or audience metrics. For contemporary analogues, use bot prevalence, engagement rates, and diffusion timestamps from platform data [6][2][8].

Case selection logic: choose cases that vary on infrastructure intensity while holding other variables (regime type, wartime exigency) constant where possible—this isolates industrialization's effect on cognitive capacity.

Causal inference: structured process tracing with focused tests of mechanism presence, strength, and causal sequence; where feasible, use interrupted time-series or regression with instrument-like indicators (e.g., introduction of railway lines, spread of telegraph connectivity) to estimate effect sizes.

Applications: parameterized vignettes, metrics, and failure modes

Vignette A — Disaster-response information campaign under intermittent communications

Context: A multinational disaster-response operation must coordinate evacuation and medical resource allocation across a region with partially disrupted networks (cell towers down in 40% of districts). The operation uses a hybrid industrialized messaging pipeline: pre-produced multilingual SMS templates (industrial-scale content), radio broadcast, and opportunistic social-media amplification via automated agents to reach mobile populations.

Parameters and metrics:

- Reach target: 80% of affected population within 48 hours.
- Mean Time to Alert (MTTA): target ≤ 4 hours from event detection to first broadcast message in affected districts.
- Message Dose: three distinct exposures within 24 hours per individual (SMS + radio + social share).
- Failure probability (systemic): modeled as $P\_failure = P\_comm\_dropout \cdot P\_misdirection \cdot P\_noncompliance$. Example numeric assumption: $P\_comm\_dropout = 0.4$, $P\_misdirection = 0.05$ (false rumor interference), $P\_noncompliance = 0.2 \rightarrow P\_failure \approx 0.4*(0.05+0.2) \approx 0.1$ (approx 10% effective failure to reach and induce compliance).

Failure modes and diagnostics:

- Infrastructure bottleneck: central dispatch cannot push SMS due to overloaded gateways. Trigger: queue length > 1000 messages/min. Delegation policy: fallback to prioritized radio broadcasts and municipal PA systems; redirection of queued messages to local cell broadcast where present.

- Adversarial/rumor interference: coordinated misinformation claiming evacuation points are unsafe. Trigger: surges in social mentions with high virality score relative to official messages (>x3). Delegation: human-in-loop verification by regional communications officer within 30 minutes; authorized counter-messaging templates issued; amplification adjusted to prioritize trusted local messengers.

- Language and legibility mismatch: templates ineffective in remote dialect communities. Trigger: low engagement/response rate in district surveys (<20% response to call-back). Delegation: local NGOs and community leaders authorized to adapt templates; A/B testing disabled until human-led validation completed.

Operational insight: industrial-scale message production reduces MTTA and permits consistent multi-channel dose, but brittle points remain (gateways, local legibility, adversarial rumors) that require human oversight and local delegation.

Vignette B — Autonomous ISR swarm with contested spectrum (cognitive influence on local populace)

Context: A non-kinetic information operation accompanies an ISR (intelligence, surveillance, reconnaissance) drone swarm that broadcasts localized PSAs (public-safety and legitimacy messaging) in contested regions. The messaging system includes automated content generation modules that personalize messages using demographic priors, all operating under contested spectrum with jamming attempts.

Parameters and metrics:

- Reach objective: 60% of target population within operational footprint within 72 hours.
- MTTA: < 6 hours for first personalized messaging pass after mission start.
- Failure probability due to jamming: P_jam = fraction of operational airspace under effective jamming; if $P\_jam = 0.3$ and failover comms reduce effective reach by 50%, resulting failure ≈ 0.15.
- Fidelity loss metric: probability that message content is truncated/corrupted during broadcast (target < 2%).

Failure modes and diagnostics:

- Contested spectrum denial: jamming raises P_jam above tolerable threshold. Trigger: signal-to-noise ratio falls below operational threshold in >10% of cells. Delegation: drones switch to store-and-forward strategy delivering printed leaflets via ground teams;

suspend personalization to avoid generating malformed or misleading fragments.

- Ethical/harm miscalculation: automated personalization generates messages that stigmatize minority groups, causing backlash. Trigger: social feedback indicating increased negative sentiment in target micro-community (>20% relative increase). Delegation policy: immediate human-in-loop review; rollback to pre-vetted, non-personalized templates; retrain personalization module only under human oversight with bias audits.

- Consensus fracture: local opinion leaders issue counter-narratives that create echo-chamber reinforcement. Trigger: rising local re-share rate of counter-narratives crossing cascade thresholds. Delegation: engage local intermediaries and trusted nodes; prioritize narrative repair via credible messengers rather than volume amplification.

Failure diagnosis and mitigation: combine automated monitoring of signal metrics and social feedback with pre-authorized escalation paths to human analysts; explicitly cap automated personalization until proven bias-limited in pilot tests.

(Combined commentary on vignettes)

These parameterized scenarios illustrate how industrialized messaging (pre-produced content, programmatic distribution) improves MTTA and dose control but introduces new failure modes—adversarial contamination, infrastructure fragility, and cultural legibility gaps —that require human oversight, fallback channels, and ethical governance. Quantitative metrics (MTTA, failure probability, fidelity loss) enable operational thresholds and delegation policies.

Limits & Open Questions

This section foregrounds operational assumptions, diagnostics, and bounded-rationality and adversarial communications models as explicit present assumptions rather than deferred questions. It also outlines open research questions.

Operational Assumptions & Diagnostics (presented assumptions)

1. Bounded-rationality assumption

- Statement: Agents (human or algorithmic) involved in cognitive operations have limited information, finite computational resources, and heuristic decision rules. Automated modules approximate ideal Bayesian updates but use heuristics for tractability.

- Concrete triggers (diagnostics):
- Anomalous divergence between predicted and observed engagement (>30% absolute deviation) over a 24-hour rolling window.
- Increasing error rates in classifier outputs for sentiment/intent (false positive/negative drift beyond historical baseline by >20%).

- Delegation policies:
- If divergence trigger fires, throttle automated personalization; switch to conservative, pre-vetted messaging templates and invoke human analyst review within a fixed SLA (e.g., 2 hours).
- Require periodic human-in-loop audits of automated decision rules whenever model retraining occurs or concept drift is detected.

1. Adversarial communications model

- Statement: Adversaries actively probe, contaminate, and exploit channels to reduce fidelity, create confusion, or mimic trusted actors. The communications environment is an adversarial signal-processing problem: defenders must assume that observed signals may be manipulated.

- Concrete triggers (diagnostics):
- Sudden burst of high-velocity messages with high structural similarity to official messages but differing provenance—indicator of mimicry campaigns.
- Coordinated account behaviors (bot-like synchronization metrics) exceeding historical thresholds in a target region [6].

- Delegation policies:
- On detection of mimicry or bot-propagated falsehoods, temporarily suspend automated amplification; deploy authenticated channels (cryptographic signatures, verified accounts) and human-mediated corrections.
- Establish pre-authorized strike/response templates for counter-messaging that prioritize credible local messengers rather than volume-based suppression.

Human-in-loop and adversarial models as present assumptions

- Human-in-loop is assumed necessary for (a) ethical adjudication, (b) local legibility adaptation, and (c) handling adversarial escalation that automated systems cannot reliably resolve. Automated systems perform routine, low-risk tasks under human oversight and explicit escalation thresholds.

- The adversary model assumes resourceful actors able to exploit both legacy industrial channels and modern AI tools. Defenders must assume adversaries can imitate, poison, and jam channels; detection relies on network- and content-based diagnostics with human adjudication.

Open empirical and theoretical questions (brief)

- Measurement challenge: How to validly operationalize cognitive outcomes (belief change, long-term norm shifts) beyond short-term engagement proxies?
- Continuity vs. rupture: Which industrial-era mechanisms remain explanatory under AI-driven personalization—are we observing incremental industrialization or a qualitative regime shift?
- Ethics and governance: What institutional architectures can balance legitimate information operations (public health messaging) and prevent abusive cognitive warfare?

- Model integration: How to formally integrate bounded-rational agent models with adversarial signal-processing frameworks to produce operational decision rules?

Implications for theory and policy

Theoretical implications

- Reconceptualize war: Cognitive modalities should be treated as central objects of strategic analysis where infrastructure and organizational routines are explanatory variables rather than mere vectors.
- Infrastructure–cognition feedbacks: Theorize the mutual reinforcement between information infrastructures and social cognition—how routinized messaging changes norms which in turn alter diffusion dynamics.

Policy implications

- Preparedness: Invest in measurement infrastructure (platform co-operation for circulation metrics, independent media audits) and in human-in-loop capacities to adjudicate adversarial or ambiguous cases.
- Regulation and governance: Foster standards for provenance (message authentication), transparency for algorithmic personalization, and cross-border norms limiting the weaponization of civilian platforms.
- Resilience and education: Prioritize civic literacy and critical-media curricula to raise population-level thresholds for cascade contagion.

Synthesis

Industrialization transformed influence by converting persuasion into an engineered, repeatable system: lowered marginal costs, increased coordination capacity, and routinized feedback—conditions that changed the scale, persistence, and organizational embedding of cognitive operations. Contemporary AI-driven systems represent a further industrialization layer: automation of production, micro-personalization, and closed-loop optimization. The combined theory suggests continuity (mechanisms remain: capacity, standardization, legibility) and acceleration (automation intensifies dose and personalization), producing both novel opportunities (rapid public-health messaging) and systemic risks (closed-loop adversarial optimization).

Operationally, resilience demands three pillars: measurement and diagnostics to detect divergence and attack; human-in-loop governance for ethical and contextual adjudication; and infrastructural redundancy to mitigate single-point failures. Research should prioritize rigorous measurement of cognitive outcomes, longitudinal studies of infrastructure–cognition feedbacks, and formal integration of bounded-rational, adversarial models.

Conclusion and avenues for future research

Contributions: This brief refines the concept of cognitive wars, links industrial-era material transformations to cognitive modalities of strategic influence, and outlines mechanisms and operational implications for modern AI-enabled systems. It provides a theory-first template for comparative-historical and contemporary empirical work.

Future research paths:

- Longitudinal case studies tracing infrastructure introduction (telegraph, radio, internet) to measurable cognitive outcomes.
- Cross-regional comparisons to identify contingent conditions (cultural legibility, institutional robustness) mediating industrial effects.
- Modeling work that integrates diffusion thresholds, bounded-rationality, and adversarial signal models to produce operational decision rules and diagnostic triggers.
- Policy experiments testing provenance/authentication tools, transparency mandates for algorithmic personalization, and civic-resilience interventions.

Footnote citations (by source id used in-text): [1] [2] [3] [4] [5] [6] [7] [8] [9] [10] [11]

Assumptions Ledger

Assumption	Rationale	Observable	Trigger	Fallback/Delegation	Scope
Industrialization (transport, mass media, bureaucratic routines, mass schooling) was a structural cause that materially enabled large-scale, durable cognitive influence operations (i.e., 'cognitive wars').	Historical changes in infrastructure and institutions—rail, telegraph, print, broadcast, mass schooling and bureaucracy—clearly lowered costs, increased throughput and coordination capacity; these changes plausibly map onto greater reach, repeatability, and persistence of messaging.	Correlations and process-traced sequences linking expansion of communication/transport metrics (print runs, telegraph traffic, newspaper circulation, literacy rates, formation of propaganda/bureaucratic bodies) to documented large-scale influence campaigns and measurable shifts in public belief or mobilization.	Encountering counter-evidence such as repeated large-scale, durable influence campaigns in pre-industrial settings, inability to find temporal/causal order between infrastructure growth and influence outcomes in case studies, or failure of historical proxies to correlate with changes in population-level cognition.	If the structural causal claim fails, narrow the thesis to a conditional claim (industrialization as an enabler in specific contexts), increase reliance on case-centered process tracing, or delegate to specialized historians/archivists to re-evaluate temporal sequences and contextual factors.	Applies primarily to industrial-era and mass-mediated societies where centralized production and mass education existed; does not necessarily generalize to pre-industrial societies, small-scale communities, or contemporary contexts with radically different socio-technical baselines without re-calibration.
Contemporary AI systems constitute a further layer of industrialization of influence by enabling automation, personalization at scale, and closed-loop optimization that materially increases influence effectiveness.	AI/ML reduces marginal production cost of content, enables targeting and personalization from behavioral data, and supports iterative optimization (A/B testing, reinforcement) that was previously impractical at scale—features documented in contemporary computational-propaganda studies and platform engineering.	Empirical signs such as high-volume personalized content generation, automated account orchestration, measurable performance improvements attributable to model-driven targeting (higher engagement, conversion, cascade initiation), and existence of closed-loop pipelines (data → model → content → measurement → retraining).	When AI-driven interventions do not outperform human-coordinated campaigns, when personalization and closed-loop behaviors are absent, or when access to training data/compute is restricted, prompting re-evaluation of the AI-as-industrializing-force claim.	Recast AI as an amplifying tool rather than a structural transformation; focus on specific socio-technical constraints (data availability, platform governance) and delegate empirical validation and ML auditing to computational social scientists and ML engineers.	Valid in environments with abundant digital traces, platform access, and computational resources; less applicable where connectivity is low, data are sparse, platforms are segregated, or robust regulation and moderation limit automated amplification.

Assumption	Rationale	Observable	Trigger	Fallback/Delegation	Scope
A theory-first, mechanism-focused methodology (process tracing + triangulation with quantitative proxies) is sufficient to identify causal links from material infrastructures to cognitive outcomes and to generate predictive leverage.	Mechanism-centered inference clarifies causal pathways and avoids conflation of heterogeneous phenomena; when paired with measurable proxies it permits falsifiable claims and bounded generalization without over-reliance on correlational inference.	Convergence between qualitative process-tracing narratives and quantitative proxies (e.g., temporal alignment between increases in channel capacity and measurable shifts in diffusion metrics), reproducibility of mechanism accounts across multiple cases, and successful out-of-sample predictions or retrodictions.	When proxies are inconsistent, data are insufficient to establish mechanism temporality, when confounders dominate correlations, or when process traces cannot be independently corroborated—prompting reassessment of methodological adequacy.	If the approach proves insufficient, expand empirical methods (archival research, experiments, surveys), drop stronger causal claims in favor of descriptive or contingent accounts, or assign empirical validation tasks to field researchers and statisticians for stronger identification strategies.	Effective when archival/proxy data exist and mechanisms are traceable; limited where data are absent, interventions are covert/deniable, or when multiple competing mechanisms produce indistinguishable outcomes without experimental leverage.
Operational metrics proposed (reach, dose, fidelity, institutional embedding, feedback) are valid, measurable constructs that meaningfully capture the dimensions of cognitive operations and can be used to assess influence campaigns.	These metrics map directly to theoretical constructs (who is addressable, intensity of exposure, message consistency, organizational durability, and adaptation loops) and correspond to measurable platform/historical indicators used in diffusion and media studies.	Successful operationalization in datasets (e.g., audience size and overlap for reach; frequency and duration for dose; content similarity measures for fidelity; organizational records for embedding; iteration logs for feedback) and statistical associations between these metrics and measured cognitive/behavioral outcomes.	When metrics cannot be operationalized with available data, when they do not correlate with cognitive changes in empirical tests, or when measurement noise overwhelms signal—indicating a need to revisit metric validity.	If metrics fail, broaden the metric set or develop new proxies (surveys, ethnography, network sampling), employ qualitative indicators, or involve measurement scientists and platform engineers to design better instrumentation and data collection protocols.	Suitable for mediated, mass-scale influence contexts and platform-mediated environments; less reliable for informal interpersonal influence, emergent social movements without clear messaging artifacts, or contexts where data access is blocked.
Canonical formal models from network diffusion, information theory, and statistical learning can be transferred and adapted to explain and	These formal models capture general properties of information transmission, capacity limits, threshold behavior, and learning/optimization dynamics that underlie how	Empirical fit: model-derived predictions (e.g., threshold-driven cascades, channel-capacity constraints, learning curves) align with observed diffusion patterns, campaign dynamics, or optimization behavior in historical and	When model predictions systematically diverge from observed realities (e.g., cultural moderation of exposure effects, non-independent	If canonical models fail, adapt them by adding domain-specific parameters, integrate qualitative socio-cultural variables, or switch to alternative frameworks (e.g., narrative/rhetorical analysis); delegate formal model	Useful at an abstract/theoretical level and for generating hypotheses; limited for micro-level or culturally-specific dynamics without careful calibration and incorporation of contextual

Assumption	Rationale	Observable	Trigger	Fallback/Delegation	Scope
predict dynamics of cognitive warfare.	messages propagate and how automated systems adapt—providing portable conceptual machinery for causal mapping.	contemporary cases after reasonable contextual calibration.	exposures invalidate model assumptions), or when key model assumptions (stationarity, independence) do not hold in the data.	adaptation to interdisciplinary teams combining modelers and area experts.	constraints and empirical heterogeneity.

Notation

Symbol	Meaning	Units / Domain
n	number of agents	$\mathbb{N}$
$G_t=(V,E_t)$	time-varying communication/interaction graph	—
$\lambda_2(G)$	algebraic connectivity (Fiedler value)	—
p	mean packet-delivery / link reliability	[0,1]
τ	latency / blackout duration	time
λ	task arrival rate	1/time
e	enforceability / command compliance	[0,1]
τ_{deleg}	delegation threshold	[0,1]
MTTA	mean time-to-assignment/action	time
P_{fail}	deadline-miss probability	[0,1]

Claim-Evidence-Method (CEM) Grid

Claim (C)	Evidence (E)	Method (M)	Status	Risk	TestID
Industrialization converted influence from episodic persuasion into an industrialized, repeatable, scalable system (a structural reconfiguration) that enabled large-scale, durable "cognitive wars".	[2] [6] [5]	Comparative-historical process tracing + archival quantification (print runs, telegraph/broadcast capacity, bureaucratic records) combined with illustrative contemporary case studies of coordinated online campaigns; causal inference via within-case mechanism tracing.	E cited; M pending archival process-tracing and cross-case synthesis	If false, the theoretical framing (structural industrialization → cognitive wars) is undermined; policy and resilience measures premised on structural, institutional levers (rather than tactical platform mitigations) may be misallocated.	T1
Industrial-era infrastructures increased reach, reduced marginal cost per exposure, and standardized fidelity—permitting sustained campaigns with measurable population-level cognitive effects (reach ↑, cost/exposure ↓, fidelity ↑ → sustained influence).	[9] [10] [6] [8]	Empirical measurement using proxies (circulation/print-run data, telegraph/broadcast throughput, platform impressions), cost-per-exposure modeling, and diffusion simulations to map cost and fidelity to achieved reach and persistence.	E cited; M pending empirical proxy analysis, cost modeling, and simulation calibration	If wrong, interventions based on reducing marginal cost or altering fidelity assumptions (e.g., platform throttling, counter-messaging investment) may not produce expected reductions in population-level influence.	T2
Contemporary AI systems constitute a further industrialization layer by enabling automation of production, micro-personalization at scale, and closed-loop optimization (automated generation + targeting + feedback → increased dose & personalization).	[3] [4] [2] [5]	System-level audits and experiments (A/B tests), controlled simulations of personalization pipelines, and field measurements of differential persuasion rates from AI-personalized vs. non-personalized content.	E cited; M pending experiments, system audits, and field validation	If false, regulatory focus on AI-specific interventions (e.g., model governance, personalization limits) could be less effective than anticipated; defensive strategies may need reorientation toward non-AI industrial practices or social remedies.	T3
Network diffusion and threshold dynamics mediate how infrastructure changes map onto cascades: increased connectivity and platform centralization lower effective cascade thresholds and increase the probability/size of large-scale cascades (i.e., diffusion sensitivity ↑ with connectivity/weak ties).	[9] [10] [1]	Agent-based and analytical diffusion simulations calibrated on empirical network/topology data; sensitivity analysis of threshold parameters and algebraic connectivity (spectral measures) to cascade probability and speed.	E cited; M pending simulation calibration and analytic sensitivity studies	If false, models predicting cascade likelihood from structural connectivity will be unreliable; mitigation strategies aimed at network structure (e.g., de-centralization, rewiring) may not reduce large-scale spread as expected.	T4

Claim (C)	Evidence (E)	Method (M)	Status	Risk	TestID
Information-theoretic channel changes (higher capacity, lower per-bit cost, reduced latency) enable higher-volume persuasive signaling: increased channel capacity raises effective throughput of influence operations and relaxes previous signal-to-noise constraints on coordinated persuasion.	[8] [5]	Formal information-theoretic modeling linking channel capacity and noise to achievable persuasive throughput; empirical estimation of message bit-rates and effective exposure rates across historical and digital channels.	E cited; M pending formal modeling and empirical throughput estimation	If wrong, arguments that target channel-level interventions (throttling volume, altering platform throughput) will reduce influence effectiveness may be misplaced; resource allocation for bandwidth or platform regulation could be suboptimal.	T5
Institutional embedding (mass schooling, bureaucratic routines, standardized media production) increases persistence and fidelity of belief-change—i.e., industrialized institutions create durable cognitive regimes that sustain influence over time.	[11] [7] [8]	Longitudinal and archival analyses linking institutional variables (literacy rates, schooling expansion, state media bureaucracies) to persistence of propagated beliefs; natural experiments where institutional shifts occurred (policy or technology shocks) to observe changes in belief durability.	E cited; M pending longitudinal archival and quasi-experimental studies	If false, resilience strategies that focus on institutional reform (education curricula, media governance) may not produce expected long-term reductions in vulnerability to cognitive warfare.	T6

Sources

[1]

On graph theoretic results underlying the analysis of consensus in multi-agent systems

Arxiv.Org, 2009-02-24. (cred: 0.50)

http://arxiv.org/abs/0902.4218v1

[2]

Online Influence Campaigns: Strategies and Vulnerabilities

Arxiv.Org, 2024-12-18. (cred: 0.50)

http://arxiv.org/abs/2501.10387v1

[3]

MetaTroll: Few-shot Detection of State-Sponsored Trolls with Transformer Adapters

Arxiv.Org, 2023-03-13. (cred: 0.50)

http://arxiv.org/abs/2303.07354v1

[4]

The Language of Influence: Sentiment, Emotion, and Hate Speech in State Sponsored Influence Operations

Arxiv.Org, 2025-05-12. (cred: 0.50)

http://arxiv.org/abs/2505.07212v3

[5]

Strategic Discourse Assessment: The Crooked Path to Innocence

Arxiv.Org, 2025-06-01. (cred: 0.50)

http://arxiv.org/abs/2506.01195v2

[6]

The Rise of Social Bots

Arxiv.Org, 2014-07-19. (cred: 0.50)

http://arxiv.org/abs/1407.5225v4

[7]

Mechanisms of True and False Rumor Sharing in Social Media: Collective Intelligence or Herd Behavior?

Arxiv.Org, 2022-07-07. (cred: 0.50)

http://arxiv.org/abs/2207.03020v3

[8]

The science of fake news

Arxiv.Org, 2023-07-15. (cred: 0.50)

http://arxiv.org/abs/2307.07903v1

[9]

"The Strength of Weak Ties" Varies Across Viral Channels

Arxiv.Org, 2024-08-07. (cred: 0.50)

http://arxiv.org/abs/2408.03579v1

[10]

Exploring the complex pattern of information spreading in online blog communities

Arxiv.Org, 2015-04-02. (cred: 0.50)

http://arxiv.org/abs/1504.00495v1

[11]

Characterizing Social Imaginaries and Self-Disclosures of Dissonance in Online Conspiracy Discussion Communities

Arxiv.Org, 2021-07-21. (cred: 0.50)

http://arxiv.org/abs/2107.10204v1

Research Roadmap

- **Phase 1 (Theory)**: Formalize claims, extend proofs, validate against canonical results
- **Phase 2 (Simulation)**: Implement stress tests, sweep parameter spaces, measure convergence/scaling
- **Phase 3 (Empirical)**: Deploy in controlled environments, collect field data, validate predictions
- **Phase 4 (Integration)**: Operationalize with human-in-loop, adversarial hardening, production deployment

Confidence Methodology: Confidence = $0.3 \cdot$ SourceDiversity + $0.25 \cdot$ AnchorCoverage + $0.25 \cdot$ MethodTransparency + $0.2 \cdot$ ReplicationReadiness, where SourceDiversity reflects unique publishers & types, AnchorCoverage reflects share of primary claims with Type-1 anchors, MethodTransparency reflects CEM completeness & assumptions ledger, and ReplicationReadiness reflects sim plan & datasets/params specified.