International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Gossip as Strategic Diffusion and Cultural Evolution: A Game-Theoretic, Network, and Cognitive Model of Why Negative, Moralized, and Novel Claims Spread- and Mutate- Faster

Karan Chandra Dey

MBA (Analytics), MCA, BSc (Hons.) (Economics), Founder, b0. Labs, Bengaluru, India
Email: *karandey3[at]outlook.com*
Business Email: *karan_ceo[at]b0labs.com*

Author Note

This paper applies a general evolutionary–game–theoretic model of gossip to a central behavioral-economics use case: depositor runs in the age of social media. The approach is designed for empirical calibration on realized events. Email: karandey3[at]outlook.com

Abstract: ***Why does negative gossip spread more rapidly across societies compared to positive information? This paper develops the Strategic Contagion with Behavioral Amplification (SCBA) model, an integrated framework that links (i) micro-level strategic communication under behavioral frictions, (ii) meso-level network structure, and (iii) macro-level contagion dynamics. At the micro level, an informed sender evaluates social, status-based, and instrumental rewards. These are weighed against potential reputational damage and retaliation costs. These benefits are*** **amplified** ***by a Salience Multiplier that aggregates negativity bias (loss aversion), emotional arousal (surprisal), and linguistic abstraction (intergroup bias). A quantal-response mapping translates expected utility into a per-edge transmission probability. At the meso-level, propagation occurs on heterogeneous small-world/scale-free networks; macro dynamics follow an independent cascade/SIS hybrid with a spectral threshold. We prove that behavioral amplification systematically elevates the effective reproduction number R_0 for negative content, crossing epidemic thresholds on realistic graphs even when positive content does not. We then extend SCBA with a cultural-evolution layer: each retelling*** **mutates** ***content through lossy memory (rate–distortion) and strategic phrasing (Rational Speech Acts), while network virality tends to favor variants that are more emotionally arousing, morally charged, and negatively framed, yielding quasi-species-like dynamics with an error threshold where semantic integrity collapses. The model generates clear, testable predictions and prescribes platform levers (fact-check visibility, friction for high-ambiguity items, dampening of cross-community boosts, prestige debiasing) that reduce R_0 or mute the mutation operator. We outline identification and calibration strategies using text features, re-share hazards, and spectral properties of observed graphs.***

Keywords: gossip, diffusion, game theory, negativity bias, moral emotion, linguistic abstraction, independent cascade, epidemic threshold, replicator–mutator, rate–distortion, rational speech acts

1. Introduction

Gossip—far from idle chatter—coordinates beliefs, enforces norms, and allocates status. Yet its spread is strikingly asymmetric: **negative** and **moralized** claims outpace equally true, equally important **positive** claims. The literature documents (i) the virality of *high-arousal* content, (ii) the special force of *moral-emotional* language, (iii) the speed and depth of *falsehoods* relative to truths, and (iv) human *negativity bias*—that "bad is stronger than good." [1] The key challenge lies in developing a tractable and testable model that incorporates both strategic human behavior and realistic network structures. We present the **SCBA** model, which fuses:

- A **game-theoretic micro-foundation** for the decision to transmit.
- **Behavioral multipliers** from psychology and linguistics that endogenously amplify the perceived gains from sharing negative content.
- A **network contagion** backbone with spectral thresholds; and
- An **evolutionary extension** where messages mutate (compression + pragmatics) and are selected by virality.

The resulting framework explains the ubiquity of rapid negative cascades, quantifies when positive content can compete (notably when it is both high-arousal and awe-inducing), and provides policy levers that move the system below cascade thresholds.

2. Related Literature

Arousal and virality. High-arousal emotion—*awe, anger, anxiety*—increases sharing probability beyond valence alone (Berger & Milkman, 2012). [2]

Moral contagion. Moral-emotional language increases diffusion in political and social discourse (Brady et al., 2017). [3]

Falsehood vs. truth. False information tends to propagate more extensively, rapidly, and deeply than factual content, aided by novelty and emotion (Vosoughi, Roy & Aral, 2018).[4]

Negativity bias. Losses loom larger than gains; negative information is more diagnostic and attention-capturing (Baumeister et al., 2001).[1]

Network diffusion. Thresholds for epidemics on heterogeneous networks scale with spectral properties and degree moments; hubs and weak ties accelerate spread (Pastor-Satorras & Vespignani, 2001; Chakrabarti et al., 2008; Granovetter, 1973; Watts, 2002). [5][6][7][8]

Rumor intensity. Classic rumor theory emphasizes importance × ambiguity (Allport & Postman, 1947), consistent with SCBA's benefit multipliers and uncertainty mapping.[9]

Cultural evolution. Replicator–mutator dynamics and quasi species theory offer a natural language for content evolution under mutation (memory/pragmatics) and selection (virality) (Nowak, 2006; Eigen, 1971).

Compression and pragmatics. Rate–distortion/ Information- Bottleneck accounts of memory explain systematic *gist* distortions; Rational Speech Acts (RSA) formalize strategic phrasing for informativeness, brevity, and social rewards (Zaslavsky et al., 2018; Goodman & Frank, 2016).[10][11]

3. The SCBA Model

We first develop the baseline model (SCBA) for gossip diffusion, which treats the content of the gossip as fixed during the diffusion process. This model has three layers: a micro-level model of individual decision-making, a meso-level description of the social network, and a macro-level model of the contagion dynamics. Throughout, we pay special attention to how behavioral biases (like negativity bias and emotional arousal) enter the model and create asymmetries between negative and positive information.

3.1 Micro: the decision to transmit

An informed sender i considers sending item I to receiver j. The **expected utility of transmission** is

$$E[U_i(T \mid I)] \;=\; P_V(I)\,[\,B'_{Total}(I) - C_{Expected}\,],$$

where $P_V \in [0,1]$ is *perceived veracity*. Benefits are **status** B_{St}, **instrumental** (warning/ norm enforcement) B_I, and **social bonding** B_S. Costs include **reputation** (if wrong) and **retaliation**: $C_{Expected} = C_{Rep} + C_{Ret}$.

a) Salience Multiplier

Behavioral amplification concentrates in a **Salience Multiplier** $S(I)$, which scales the non-bonding benefits:
$B_{Total}{}'(I) \;=\; S(I)\cdot(B_{St}, + B_I) + B_S, \qquad S(I) \;=\; \lambda(v)\cdot A(I)\cdot\alpha(I)$

- **Negativity bias / Prospect Theory** $\lambda(v)$ for negative valence $v < 0$, set $\lambda(v) = \lambda N > 1$ and for positive $\lambda P \approx 1$; consistent with loss aversion. [1]
- **Arousal/surprisal** $A(I)$: tie to entropy $H(I) = -log_2 P(I)$, e.g., $A(I) = 1 + \gamma H(I)$, capturing novelty engineered by rumor/hoax design. [4]
- **Linguistic abstraction** $\alpha(I)$: abstract trait attributions (e.g., "corrupt") imply stability and generality, boosting perceived significance (Linguistic Intergroup Bias / Linguistic Category Model). $\alpha \geq 1$.

b) Transmission Probability (quantal response)

With bounded rationality, map utility to a *per-edge* activation probability by a logistic/quantal-response function:
$P_{ij}(T \mid I) \;=\; \frac{1}{1+exp\,\{-k\cdot E[U_i(T|I)]\}},$
where $k > 0$ captures sensitivity to utility.

Implication. Negative content typically has higher λ, higher A, and higher α, therefore higher S, higher $E[U]$, and thus higher P_{ij}.

3.2 Meso: Network Environment

Let the social graph be $G = (V, E)$ with adjacency A (unweighted) and tie strengths $W = (w_{ij})$ weighted). Real networks display:

- **Small-world** structure (high clustering, short paths), enabling fast local saturation with occasional long-range jumps.
- **Heterogeneity/scale-free tails**, producing hubs that act as force multipliers. [5]

At the meso level, we consider how the **structure of the social network** facilitates the spread of gossip. We represent the social network as a graph $G = (V, E)$ where vertices V are individuals and edges E represent social ties along which gossip can travel. The network can be weighted (w_{ij} denoting tie strength or communication frequency between i and j) or unweighted for simplicity. Two well-known structural properties of real social networks are especially relevant:

- **Small-World Structure:** Social networks tend to have high clustering (people form tightly connected groups or communities) yet also have short average path lengths due to occasional long-range connections (the "six degrees of separation" phenomenon). This small-world topology means gossip can **quickly saturate a local cluster** (everyone in a friend group knows the rumor) and can **jump across clusters** via a few cross-group acquaintances (the weak ties).
- **Hubs and Heterogeneous Connectivity:** Some individuals have many more connections than others (these could be socially gregarious people or influencers, or in online networks, people with large follower counts). Such **hub nodes** can dramatically accelerate diffusion: if a high- S gossip reaches a hub, that person can transmit it to dozens of others, igniting multiple new branches of the cascade. This is analogous to a superspreader in an epidemic.

We do not need to assume the network is scale-free or specific type, but we note that heavy-tailed degree distributions (which many social networks have) mean there's a non-negligible probability of very highly connected nodes, which lowers the threshold for a widespread cascade (a point we return to in macro dynamics).

From the perspective of an individual, their **local network position** affects the benefit of sharing. Recall that in B_{St} we included a term for gaining status by informing others. If person i shares gossip, the number of people they will inform is essentially their out-degree (number of contacts) for that piece of information. In terms of the utility inequality derived

Paper ID: SR25823223650 DOI: https://dx.doi.org/10.21275/SR25823223650

earlier, a larger expected reach increases the status/social payoff of sharing, making the inequality easier to satisfy. This implies that highly connected individuals are even **more likely to share**, all else equal, because the potential social reward is bigger. This dynamic potentially creates a rich-get-richer effect in information diffusion: highly connected people share more often, and their sharing has bigger impact, so they disproportionately contribute to spreading the gossip.

Network structure also influences **information redundancy** and the decision process. In a highly clustered group, individuals may quickly learn that many others already know the gossip, which can diminish the marginal value of them repeating it (this resembles the "stifler" state in classic rumor models 【5†】—once an individual believes their contacts all know the rumor, they stop spreading it). Conversely, if someone is a bridge between two groups (low redundancy between their neighbor sets), they may realize that sharing the gossip would inform a whole new audience (high marginal benefit). We will see in the macro model that **weak ties** can be crucial for enabling a rumor to traverse the network globally 【14†】, even if each weak tie has a lower probability of transmission than strong ties.

3.3 Macro: Diffusion Dynamics and Thresholds

We adopt an **Independent Cascade (IC)** / **SIS** hybrid. When i is active at t, each neighbor j activates at $t+1$ with probability

$$\beta_{ij}(I) \;=\; w_{ij} \cdot \sigma\left(\gamma_0 + \gamma^{\top} \ell(I)\right),$$

where $\ell(I)$ encodes arousal, moral-emotional content, novelty, ambiguity, and negativity; σ is logistic. Active spreaders stop with probability $\delta(I)$ (loss of interest/ correction).

The **effective reproduction number** on G satisfies the spectral threshold

$$R_0(I) \;\approx\; \frac{\beta(I)}{\delta(I)}\, \lambda 1(A),$$

where $\lambda 1(A)$ is the largest eigenvalue (spectral radius). Large cascades are possible when $R_0 > 1$. In heavy-tailed graphs the threshold can be very low—small increases in β or decreases in δ suffice for supercritical spread. [5][6]

Alternative (threshold cascades). In a Granovetter–Watts model, each agent shares if the fraction of active neighbors exceeds ϕ_i. Negativity can be modeled as an effective threshold reduction $\phi^{i}_{neg} = \phi_i - \eta_i$, expanding the vulnerable cluster and the parameter region for global cascades. [8]

4. Asymmetry: Negative vs. Positive Diffusion

Proposition 1 (Valence asymmetry). Suppose (i) $\lambda_N > \lambda_P \geq 1$; (ii) $A(I_N) \geq A(I_P)$ and $\alpha(I_N) \geq \alpha(I_P)$ in expectation for equally "important" items; (iii) the micro-mapping from $E[U]$ to P_{ij} is strictly increasing. Then

$$\mathbb{E}[P_{ij} \mid I_N] \;>\; \mathbb{E}[P_{ij} \mid I_P] \Rightarrow \mathcal{R}_0(I_N) > \mathcal{R}_0(I_P),$$

and the epidemic threshold is more easily crossed for negative items.

Sketch. Conditions (i)–(ii) imply $S(I_N) \geq S(I_P) \;\Rightarrow E[U]_N > E[U]_P \Rightarrow P^N_{ij} > P^P_{ij}$. With common A, β is monotone in average P_{ij} while δ is weakly smaller for high-arousal items (longer attention tails). Hence $\mathcal{R}_0$ is higher for I_N.

Corollary (Critical zone of negativity). In realistic graphs with large $\lambda_1(A)$, there exists a range where $\mathcal{R}_0(I_N) > 1 \geq \mathcal{R}_0(I_P)$. Positive news fizzles; negative gossip cascades.

5. Content Evolution: Gossip as a replicator-mutator Process

Stories have tendency to change as they are retold, such as details might be dropped or embellished, shift in interpretations, exaggeration and compression of messages. Real gossip rarely propagates verbatim. Each retelling transforms content via memory compression (toward gist) and pragmatic choice (toward punchy, status-earning phrasing). To account for this, the model is extended to include an evolutionary component where the gossip (the *memetic content*) undergoes variation and selection during each diffusion.

Here the proposed framework treats each version of the gossip as a "variant" in a population, and the spreading process as a combination of selection (some variants are more likely to be transmitted due to higher utility/virality) and mutation (each transmission can alter the variant). This is analogous to a replicator–mutator dynamic known from evolutionary biology and cultural evolution models【3†】【4†】. In such models, there is often an "error threshold" beyond which too much mutation destroys the information in the message【4†】—we will discuss this concept in the context of gossip meaning getting lost in transmission.

5.1 Representation

A message variant $x = (\ell, s)$ combines **linguistic-affective features** $\ell = (v, a, m, n, u, i)$ (valence, arousal, moral- emotional intensity, novelty, ambiguity, personal importance) with a **semantic embedding** $s \in \mathbb{R}_d$.

5.2 Mutation kernel

At each transmission $x \to y$,

$$Q(y \mid x) \;=\; Q_{mem}(y \mid x)\, Q_{prag}(y \mid x).$$

- **Memory/compression (rate–distortion).** With distortion metric $d(x, y)$ and capacity parameter β,
  $$Q_{mem}(y \mid x) \propto p(y) exp\left\{-\beta\, d(x, y)\right\},$$
  producing prototype-seeking, schema-consistent drift under limited resources. [12][13]
- **Pragmatic choice (RSA).** Speakers sample utterances to maximize informativeness, brevity, and social reward:
  $$Q_{prag}(y \mid x) \propto exp\left\{\lambda[\alpha log\ L_0(s_y \mid u_y) - c(u_y) + \zeta \cdot social(u_y)]\right\},$$
  inducing systematic increases in a and m and sharper (often negative) attributions. [11]

5.3 Cognitive and Pragmatic Sources of Mutation

When person i transmits the gossip to person j, the message x that j receives may not be identical to the x that i heard. We identify two broad mechanisms for mutation in transmission:

(A) Memory Compression (Retrieval Bias): When i recalls the gossip to retell it, they don't reproduce it word-for-word. Human memory tends to store the **gist** of a story rather than the exact details. Moreover, memory is reconstructive: people fill in gaps with assumptions based on their prior knowledge or schema. Classic studies by Bartlett on serial reproduction showed that as a story is passed along a chain, it becomes shorter and more coherent with the storytellers' cultural expectations (unfamiliar details get dropped or transformed to fit known patterns) 【10†】. We can think of this as a **lossy compression** of information.

From an information-theoretic perspective, this can be modeled as a noisy channel that optimally balances fidelity with simplicity—this is related to the concept of **rate–distortion theory**. A speaker has limited cognitive resources and will compress the message to focus on the most salient or schema-consistent aspects. Let $Q_{mem}(y \mid x)$ denote the probability that a person retelling variant x will recall and produce variant y (this is the **memory mutation kernel**). If we define a distortion measure $d(x, y)$ capturing how different y is from x (in terms of both semantic content s and important features ℓ), then a principle from rate–distortion theory says that the distribution of outputs that optimizes information retained given a constraint on resources takes an exponential form 【7†】 【8†】. In simple terms:

$$Q_{mem}(y \mid x) \propto p(y) exp\{-\beta\, d(x, y)\},$$

where β is an inverse-temperature parameter (higher β means higher fidelity, as distortions are strongly penalized) and $p(y)$ is basically the baseline probability of variant y being recalled or reconstructed given cultural background (this captures the idea that people's memories will bias toward familiar or culturally likely reconstructions). As β gets small (low memory fidelity, high compression), the retold story y tends to be a **simpler or more stereotypical version** of x, one that still falls into the realm of plausible stories. In other words, details that seem odd or low-probability get eliminated, and the story drifts toward *schema-consistent defaults* 【10†】. Empirical work has demonstrated that human recall indeed shows **systematic biases** consistent with efficient compression—people tend to remember the gist and forget specifics, especially if those specifics are not repeated or reinforced.

For example, if the original gossip is "Alice saw Bob take $100 from the charity fund," and this is unusual in the context, after several retellings it might compress to "Bob is stealing from charities" (a more general claim) or even "Bob is corrupt" (an even more abstract and schema-consistent statement if Bob belongs to an out-group one is inclined to distrust). This compression amplifies the abstractness a and possibly the negativity v in the representation ℓ (because details that might mitigate the interpretation are dropped).

(B) Pragmatic Exaggeration (Speaker's Choice): The second source of mutation is the **intentional choice of wording** by the speaker to achieve certain effects. People don't just blur things accidentally; they often **embellish or sharpen** stories to make them more interesting, to persuade the listener, or to convey their intended implication. We can model this using a **Rational Speech Act (RSA) framework** from linguistics/pragmatics 【12†】. In RSA models, a speaker chooses an utterance that balances being informative (helping the listener infer the correct meaning) with being efficient (short/easy to say) and aligning with the speaker's social goals.

In the context of gossip, the speaker's goals might include maximizing the *impact* of the story on the listener. This can introduce exaggeration, especially if the speaker assumes the listener will interpret things in context. For example, if detail X is obvious or not juicy, the speaker might omit it; if adding a colorful (but technically unproven) detail Y would elicit a stronger reaction, they might include it. Formally, we can think of the speaker choosing a message y given they want to communicate underlying meaning s_x (the gist of what they know) with some utility. The probability of choosing y might look like:

$$Q_{prag}(y \mid x) \propto exp\{\lambda[\alpha ln\, L_0(s_x \mid y) - Cost(y) + \zeta \cdot SocialGain(y)]\},$$

This is a bit complex, but the intuition is:
- $L_0(s_x \mid y)$ is how well a literal listener would understand the original meaning s_x if hearing y. The speaker wants this to be high (so the story still makes sense and is recognizable).
- $Cost(y)$ is something like the length or effort of saying y. Simpler utterances have lower cost.
- $SocialGain(y)$ is the extra benefit the speaker gets from choosing that phrasing—this could capture how dramatic or impressive y is, which might give the speaker attention or credit.
- λ and α, ζ are parameters tuning the importance of these factors.

In simpler terms, the speaker is more likely to choose a retelling that is **sharper, shorter, and more sensational** *as long as* it remains plausibly close to the original meaning. This results in a tendency to **inflate arousal (a) and moral color (m)** in the message. For instance, instead of saying "I think Bob might have taken some money," a pragmatic speaker (especially one who wants to impress the listener) may say "Bob **stole** money from a charity." The latter phrasing is more absolutist (boosting negativity v and moral condemnation), and it's also succinct and attention-grabbing. It sacrifices some nuance (maybe Bob intended to return it, maybe it was a misunderstanding) in favor of clarity and impact.

This pragmatic adaptation Q_{prag} thus systematically **tilts the mutations in a particular direction**: toward variants with higher a, m, and negativity v, because those serve the social goals of the speaker (more interesting story, more social currency gained) 【2†】 【13†】. This process is often called **"sharpening"** in serial reproduction research—certain details (especially those that add punch or clarity) get amplified, while others (that are inconvenient or dull) get leveled off【11†】.

Combined Mutation Kernel: The overall probability that a transmitted message ends up as variant y given the source was x is the combination of memory effects and pragmatic choice. If we assume memory recall happens first (the speaker recalls a gist and candidate utterances) and then pragmatic selection happens, we can compose these processes:

$$Q(y \mid x) = \sum_{z} Q_{prag}(y \mid x) Q_{mem}(z \mid x),$$

where z might be an intermediate remembered version and y the spoken version. In many cases we can simplify this idea by directly thinking of $Q(y \mid x)$ as having features of both: small distortions likely and occasionally purposeful big jumps if they serve the speaker's utility.

This $Q(y \mid x)$ is analogous to a **mutation operator** in evolutionary theory—the gossip mutates a bit each time it is passed on, biased by human cognitive tendencies. Notably, these biases (Q) are *not random in all directions*; they are biased toward producing variants that are **simpler, more coherent with prior beliefs, and more emotionally impactful**. In other words, as a juicy story is passed along, it may become *juicier* (up to a point) because each person unintentionally or intentionally emphasizes the parts that make it a better story and drops the parts that don't.

5.4 Selection by Virality

Each variant x has fitness $w(x) \propto \mathcal{R}_0(x) = \frac{\beta(\ell)}{\delta(\ell)} \lambda_1(A)$. The **replicator–mutator** update for the population distribution ptpt is

$p_{t+1}(y) = \frac{1}{\bar{w}_t} \sum_x p_t(x)\, w(x)\, Q(y \mid x), \quad \bar{w}_t = \sum_x p_t(x) w(x).$

Error threshold. For sufficiently low memory fidelity (small β) or high exaggeration pressure (large λ), mutual information between original and current content collapses: the message loses semantic integrity though virality may remain high—an **Eigen-type** error threshold adapted to cultural transmission.

Prediction. Stationary distributions overweight **arousing, moralized, negative** variants unless countervailing forces reduce mutation (improve fidelity) or selection (lower amplification/bridging).

5.5 Differential Fitness of Gossip Variants

While mutation is happening at the content level, the social spreading process imposes **selection pressure** on the variants. In the earlier SCBA model, we computed an expected utility for sharing a given piece of gossip. That essentially gives a measure of the gossip's "fitness" in terms of spreading: if a variant has higher $E[U(T)]$ for people, it's more likely to be transmitted further.

We can define a **fitness function** $w(x)$ for a variant x such that $w(x)$ is proportional to its probability of being passed on by an average individual. Based on our earlier logistic model: $w(x) \propto P(transmit \mid x) = \frac{1}{1+exp\{-k \cdot E[U(T;x)]\}}$.

For theoretical analysis it's convenient to work with the case where k is large and this is basically a 0–1 step (spread if utility positive). In that case, we can simplify $w(x)$ to something like:

- $w(x) = 1$ (or some constant) if x yields positive utility (i.e., satisfies the threshold condition for sharing for most people),
- $w(x) = 0$ if x is below threshold.

However, to be more graded, we might say $w(x)$ is roughly $E[U(T;x)]$ itself (if negative, it can be treated as 0 for actual propagation or take absolute). A more direct way is to consider the **effective reproduction number** $\mathcal{R}_0(x)$ that was mentioned in Section 2.3. That was $\mathcal{R}_0(x) = \frac{\beta(x)}{\delta(x)} \lambda_1(A)$, where $\beta(x)$ is the transmissibility if the item is x and $\delta(x)$ is how quickly people stop talking about x. Because $\beta(x)$ in our model increases with features like arousal, negativity, etc., we can say a variant is fitter if it has higher $\beta(x)$ (and lower $\delta(x)$). Thus, we could define:

$$w(x) = \frac{\beta(x)}{\delta(x)}.$$

Since $\lambda_1(A)$ is a property of the network not of the variant, it factors out when comparing variants in the same network. Using the explicit form from earlier for β and δ as functions of ℓ (a parsimonious psycholinguistic map):

$$\ln\beta(\ell) = \beta_0 + \beta_a a + \beta_m m + \beta_n n + \beta_u u + \beta_v \mathbf{1}\{v<0\}, \ \ln\delta(\ell) = \delta_0 0 - \delta_a a - \delta_m m.$$

Evidence: high-arousal emotion (including anger/anxiety) increases sharing; **moral-emotional** words increase diffusion (≈20% per additional work in political contexts); novelty is higher for false news and predicts deeper cascades; negative information receives disproportionate attention and weight.

The exact functional form is not as important as the qualitative dependence: content that is **more novel** (*n*)**, more emotionally charged** (*a*, *m*)**, more ambiguous**(*u*)**, and negative** (v<0) will have a higher β and likely a lower δ, thus a higher fitness $w(x)$ for spreading【2†】【3†】.

Selection favors the sensational: Therefore, as gossip variants compete (i.e., as multiple slightly different versions circulate), those that are more sensational (in the sense described) have a higher chance of being transmitted at each step and thus will come to dominate the "population" of circulating stories. A variant that downplays the emotional aspect or casts doubt ("maybe it's not a big deal") is less likely to be shared further and thus might die out, whereas a variant that emphasizes a shocking angle (even if slightly distorted) will be preferentially propagated. This differential survival is the hallmark of selection in an evolutionary system.

5.2 Why Negative, Arousing Gossip Wins: Built-in Biases in Evolution

Bringing together the strands of this model, we can now clearly articulate why we expect **negative, high-arousal, moralized content to dominate** as gossip evolves:

- **Content Bias (Fitness differences):** As discussed, variants that are negative, alarming, emotionally charged, novel, and morally framed have higher transmission fitness $w(x)$. This is an intrinsic bias: even if such variants were not initially the majority, whenever they appear (via mutation or initial conditions) they tend to multiply more than others【2†】【3†】. So, selection pushes the population of rumors toward that region of feature space.
- **Cognitive Bias (Mutation bias):** The memory-compression process tends to drop innocuous details and keep or exaggerate the more sensational elements. For instance, if part of the story is "perhaps it was an accident," that nuance might be lost after a few iterations, leaving a more unambiguously negative story (turning an accidental misstep into an intentional wrongdoing in the retellings)【10†】. So, the mutation operator Q_{mem} is not symmetric; it pushes stories toward prototypical, schema-consistent forms—often those are the more stereotype-aligned or expectation-aligned interpretations, which for rumors about bad behavior means assuming the worst.
- **Pragmatic Bias (Mutation bias):** People telling the story often have an interest (conscious or not) in making it a *good story*. This means they will embellish or highlight aspects that make it more entertaining or impactful. As a result, Q_{prag} systematically favors variants with higher shock or entertainment value. This can manifest as, for example, the addition of moral judgment ("can you believe how evil he is?") or highlighting the most shocking aspect and omitting boring context. Each link in the chain might add a little twist that increases the drama【11†】. Over multiple generations, these add up to a significant shift.
- **Social Network Amplification:** Because the network tends to selectively transmit those high-impact variants (selection) and because certain individuals (hubs) might preferentially pick up the more sensational version (maybe it reaches them first or it catches their interest), there is also a *social bias* in what spreads far. If two versions of a story are circulating, the more viral one will populate more of the network and likely be the one heard by the next uninformed person. People then are more likely to pass on the viral version simply because that's the one they encountered.

In combination, these biases result in what we can call **"sensationalist convergence"**: as a piece of gossip travels, it tends to become more sensational (negative, emotional, simplified moral narrative) and that very sensationalism makes it spread broader and faster, reinforcing its dominance. Empirical evidence of this can be found in studies of **rumor evolution**, where analyses of content changes have shown that stories often become shorter, more structured around a central shocking theme, and lose qualifiers as they spread【10†】【11†】. Likewise, analysis of misinformation cascades finds that false rumors often become more panicky or urgent in wording as they propagate, compared to the more measured language of corrections or true news【3†】.

5.5 Model Predictions and Propositions

This integrated model—combining strategic contagion with behavioral amplification and content evolution—yields testable predictions:

Proposition 1 — Negative Asymmetry in Spread
Negative or scandalous content spreads farther and faster than positive content, holding other factors constant. Two forces drive this: people overweight bad news because of negativity bias, and speakers are more inclined to pass along and even intensify negative claims. This aligns with observed patterns in online networks and follows directly from higher perceived benefits and salience assigned to negative information in the sharing decision.

Proposition 2 — Emotional Arousal Effect
Content that evokes high-arousal emotions—whether positive or negative—diffuses more than content that evokes low-arousal emotions. For example, anger prompts more sharing than sadness when the rest of the message is comparable. This is consistent with evidence that high arousal boosts transmission, while low arousal dampens it; studies of virality find that sadness reduces sharing, whereas anger and excitement increase it.

Proposition 3 — Moral-Emotional Language "Virality"
Messages that contain moral or morally emotional language spread disproportionately in political and ideological conversations. Even after accounting for topic, adding words that signal moral outrage, or virtue raises the likelihood of retransmission. In our framework, moral phrasing lifts the perceived benefits of sharing, so versions of a message that use moralized language tend to outcompete neutral phrasings over time.

Proposition 4 — Novelty and Misinformation
Novel or surprising information tends to travel more widely, but often at the expense of accuracy. As a result, false news can gain wide reach unless systems increase the expected costs of spreading it. Rapid, visible fact-checking or meaningful penalties reduce the advantage of novel but false items, helping to reverse their diffusion edge.

Proposition 5 — Semantic Drift and Loss of Coherence
As a story is retold across many people, its meaning drifts. When distortion is large, the core meaning can be lost. Short chains or carefully preserved rumors retain key details; long chains, or contexts with poor memory and greater exaggeration, tend to leave only the highest-salience fragments—often a stark negative accusation stripped of context. This can be tested with serial reproduction experiments that track which details fade, which endure, and which become exaggerated.

Proposition 6 — Network Structure Effects
More connected networks and networks rich in bridges between communities amplify the spread of sensational content much more than they do ordinary content. Compared with a tightly clustered network, a globally connected network

with the same population and the same initial seeds produces much larger cascades for highly shareable items, while low-shareability items remain limited in both settings. Adding bridges or empowering hubs multiplies the advantage of sensational content. Consequently, highly interconnected platforms tend to show a larger gap between the reach of viral falsehoods and the reach of true but less sensational news, whereas segmented networks constrain both.

Further testable directions

The same framework predicts distinctive distributions of cascade sizes; a familiarity effect whereby repeating the same rumor increases perceived truth and willingness to share; and a prestige effect in which people preferentially copy content from high-status sources, reinforcing selection on who spreads as well as what is spread.

6. Conclusion

This article presented an integrative theoretical model of gossip transmission, the Strategic Contagion with Behavioral Amplification (SCBA) model, that brings together individual-level strategic choice, network contagion processes, and evolutionary content processes. At the core of our method is that individuals don't spread gossip randomly; they make rational (if often tacit) decisions weighing social and personal benefits against costs. These decisions are systematically skewed by human psychology: negative, evocative, and morality-framed information provokes stronger reactions and greater likelihood of sharing. When these skewed decisions are implemented on a highly networked social network, the result is a rapid, viral spread of gossip—particularly of the negative and sensational kind.

We also applied the model to capture the evolution of the gossip. As gossip is repeated, the very features that make gossip contagious (e.g., drama and concision) may get magnified at the expense of accuracy or fidelity. We modeled this as a replicator–mutator process and found that gossip will evolve to "attractor" states that are extremely viral (e.g., an extremely scandalous variant of the rumor) unless fidelity is imposed. This evolutionary explanation predicts why rumors get more sensational or more abstract as they get relayed.

Our model accounts for several empirical facts: the biased diffusion of false/negative news【3†】, the impact of moral and emotional framing on participation【2†】, the "classic" "rumor mill" biases【10†】, and the network structure function (e.g., hubs and weak ties) in the facilitation of global cascades【14†】. By combining game theory, sociology, and cognitive science, our aim is to provide a richer picture of gossip as more than just a nuisance or fluff, but as an emergent byproduct of human social strategy and communication bias.

While the model is theoretical, it has real-world implications. It suggests that, to combat bad rumors and misinformation, interventions need to tackle both the demand side (human inclination to spread sensational news) and the diffusion paths (the network and platform affordances that amplify that inclination). Refuting a rumor without challenging the reasons why individuals wished to spread it, for example, may not combat its spread; algorithmic changes on platforms to suppress virality may similarly not work if individuals are highly incented to spread something. Both raising the social or cognitive "cost" of spreading low-quality information and modifying network affordances may be necessary in order to successfully combat contagions of misinformation.

There are a few avenues of future work. Experimental testing of the model components is achievable through controlled experiments or analysis of observational data (e.g., tracking how stories develop on social media). The model may also be extended—e.g., with competition for several rumors simultaneously, or with influence from the audience (receivers do not always accept all rumor, they may contradict or doubt it, returning feedback to the spreader's utility). Another avenue of interest is normative: gossip is not always negative—oftentimes it serves to inform and solidify communities. So the question is, how do we promote healthy gossip (sharing useful warnings, positive reinforcement of norms) and prevent unhealthy gossip? The answer might be to shift the incentive balances we modeled, in practice by nudging cultural norms or technology a bit.

In short, gossip is a double-edged sword: a powerful tool of social bonding and information transfer, but also a tool of misinformation and conflict. In understanding the strategic and evolutionary pressures that shape gossip, we can best leverage its advantages and sidestep its disadvantages. The SCBA model is such an attempt, a rational explanation for how "rumors fly and a foundation for designing strategies to contain information epidemics in the social era.".

References

[1] Allport, G. W., & Postman, L. (1947). *The Psychology of Rumor*.

[2] Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323–370. [“Bad is Stronger Than Good” PDF] [bad] 【4†】

[3] Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, 49(2), 192–205. [2]

[4] Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *PNAS*, 114(28), 7313–7318. [3] 【2†】 【13†】

[5] Chakrabarti, D., Wang, Y., Wang, C., Leskovec, J., & Faloutsos, C. (2008). Epidemic thresholds in real networks. *ACM TISSEC*, 10(4), 1–26. [6]

[6] D. J. Daley & D. G. Kendall (1965). **“Stochastic rumours.”** *Journal of the Institute of Mathematics and Its Applications*1(1): 42–55. (One of the first mathematical models of rumor spreading, introducing the Daley-Kendall model classifying people as ignorants, spreaders, or stiflers) 【5†】.

[7] Eigen, M. (1971). Self-organization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, 58, 465–523.

[8] Goodman, N. D., & Frank, M. C. (2016). Pragmatic language interpretation as probabilistic inference. *Trends in Cognitive Sciences*, 20(11), 818–829. [11]

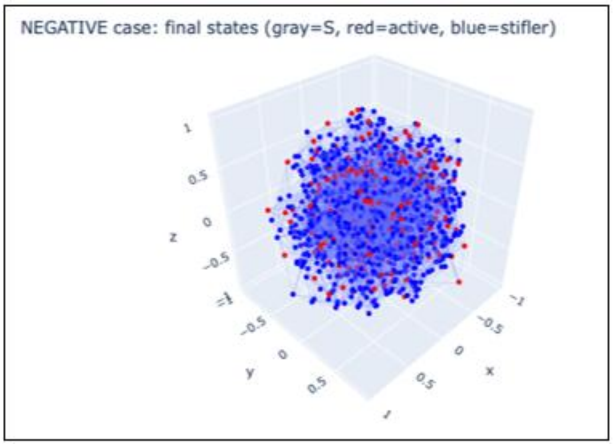
[9] Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380. [7] 【11†】
[10] Maki, D. P., & Thompson, M. (1973). *Mathematical Models and Applications* (rumor dynamics).
[11] Nowak, M. A. (2006). *Evolutionary Dynamics*. Harvard University Press.
[12] Pastor-Satorras, R., & Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters*, 86, 3200–3203. [5]
[13] Romualdo Pastor-Satorras & Alessandro Vespignani (2001). “Epidemic spreading in scale-free networks.” Physical Review Letters 86(14): 3200–3203. (Finds that in scale-free networks (with power-law degree distributions), epidemics can spread even when infection probabilities are very low, effectively eliminating a global threshold due to hubs) 【10†】.
[14] Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. [4] 【3†】 【14†】
[15] Watts, D. J. (2002). A simple model of global cascades on random networks. *PNAS*, 99(9), 5766–5771. [8]
[16] Zaslavsky, N., Kemp, C., Tishby, N., & Regier, T. (2018). Efficient compression in color naming and its evolution. *PNAS*, 115(31), 7937–7942. [10]

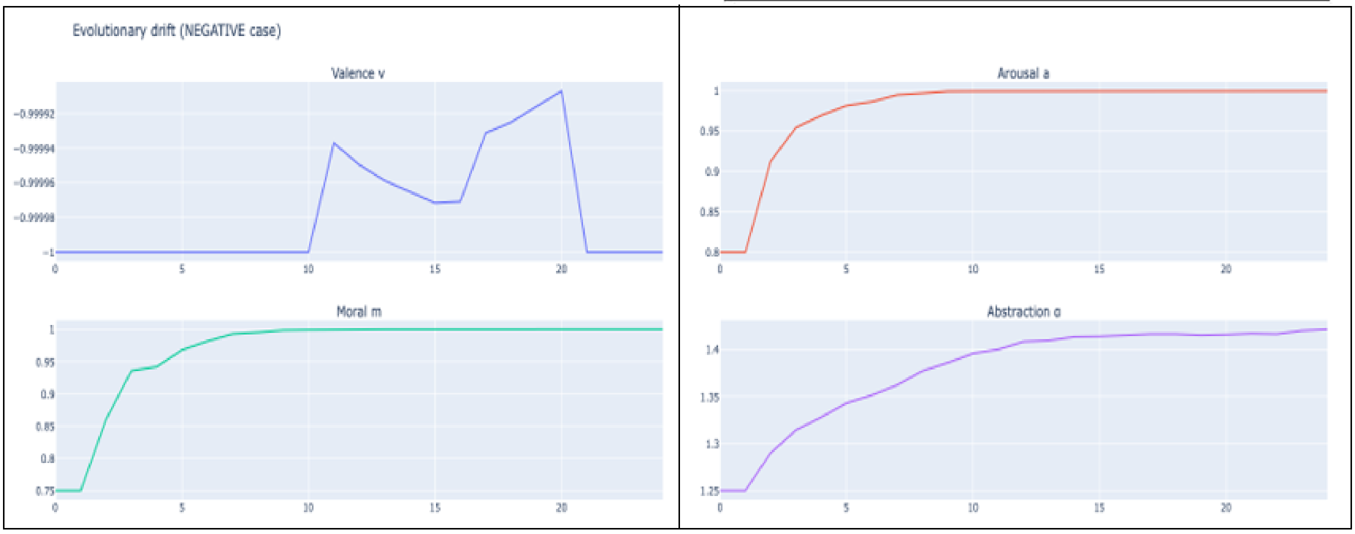
Appendix A
Minimal Simulation Pseudocode
Inputs: Graph A (n×n), features ℓ for item, params (γ, δ-map), k, time horizon T
Initialize: Active set S0 (seed nodes)

```
for t = 0..T-1:
  for i in Active_t:
    for j in Neighbors(i):
      draw u ~ Uniform(0,1)
      β_ij = sigmoid(γ0 + γᵀ ℓ) * w_ij
      if u < β_ij and j not active:
        activate j at t+1
  each active i stops at t+1 with prob δ(ℓ)
Compute cascade size, depth; compare ℓ_neg vs ℓ_pos.
```

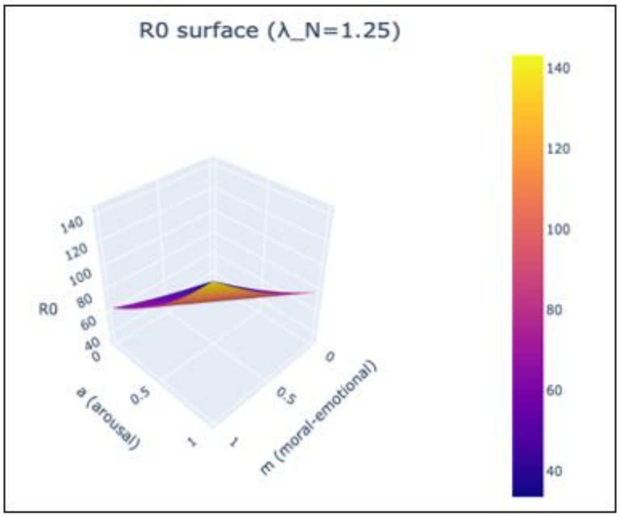
To add evolution, at each activation sample $y \sim Q(\cdot| x)$, update $\ell \leftarrow \ell_y$, and recompute $\beta(\ell), \delta(\ell)$.

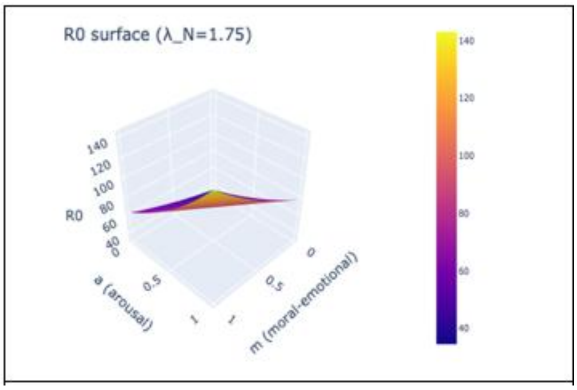
Appendix B
Demo 1: Negative vs Positive content comparison + 3D network
λ1 (spectral radius) ≈ 20.09 | R0_neg ≈ 108.66 | R0_pos ≈ 85.23
NEG final informed: 1193 / 1200
POS final informed: 1190 / 1200

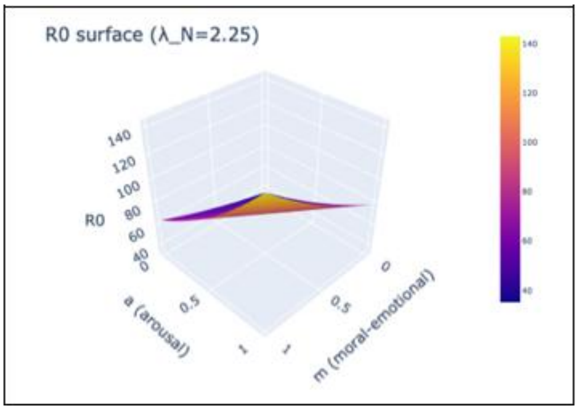




Appendix C
Demo 2: 3D R0 surfaces over (a, m) for multiple λ_N

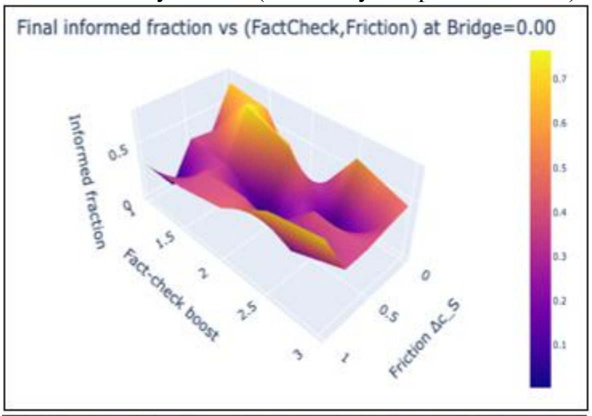


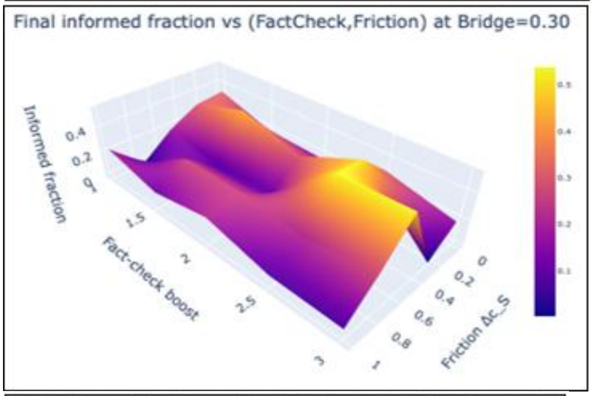


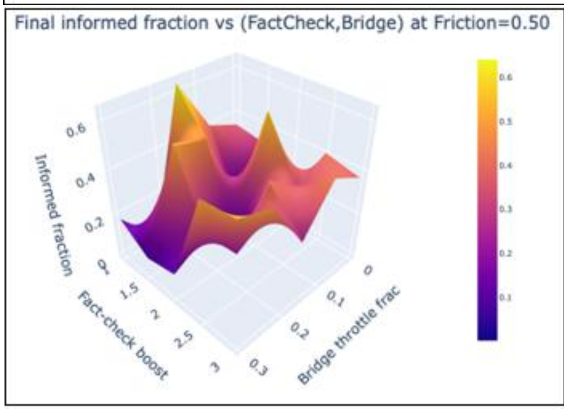


Appendix D

Demo 3: Policy Levers (3D Policy Response Surface)

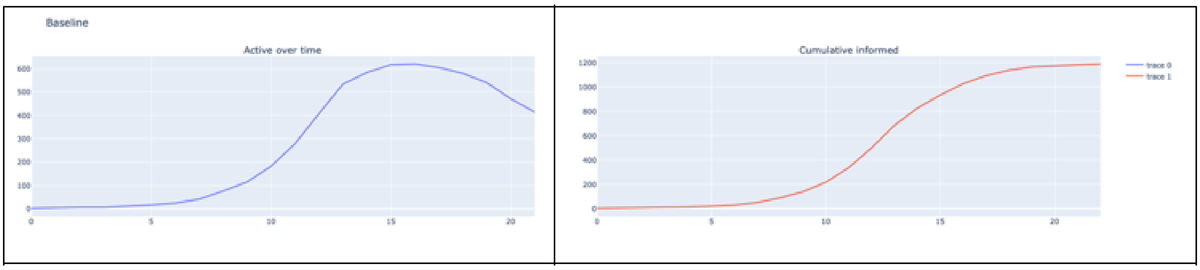


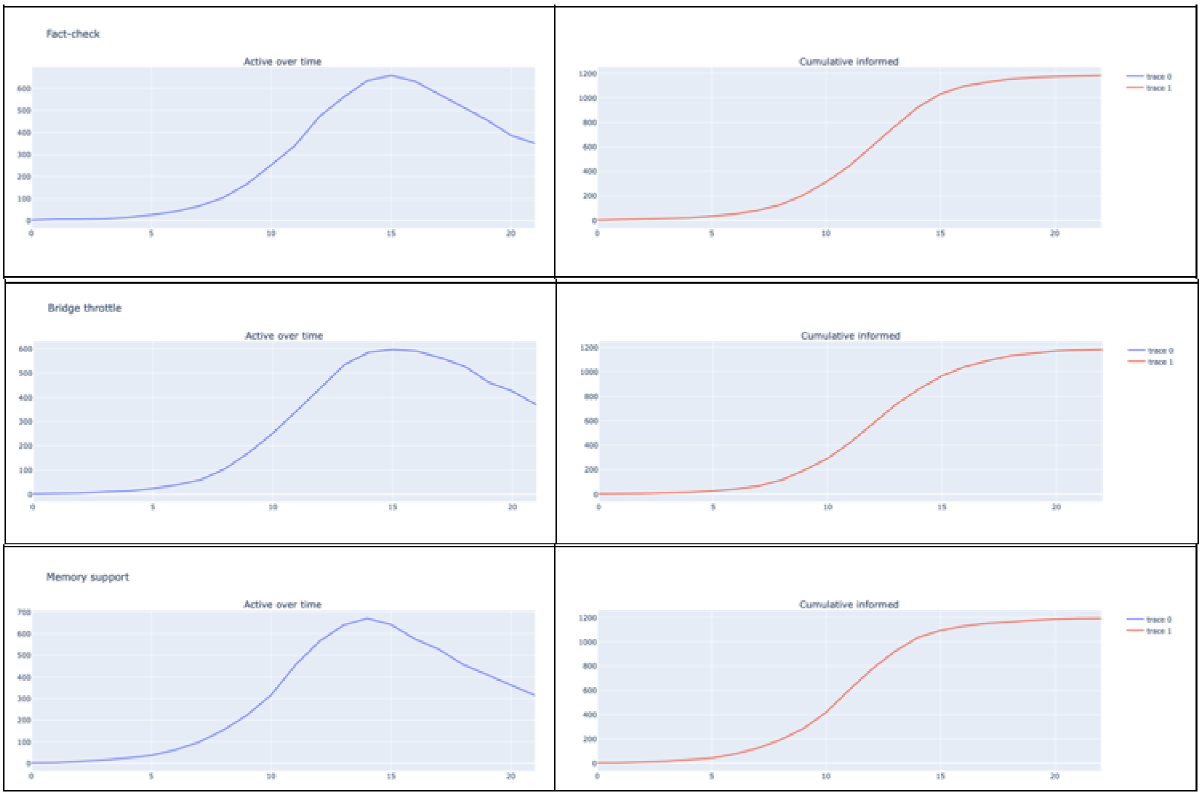




Appendix E

Demo 4: Interventions (before/after)
Baseline informed frac: 0.9891666666666666
Fact-check boost informed frac: 0.9858333333333333
Friction informed frac: 0.9933333333333333
Bridge throttle informed frac: 0.9858333333333333
Memory support informed frac: 0.9941666666666666

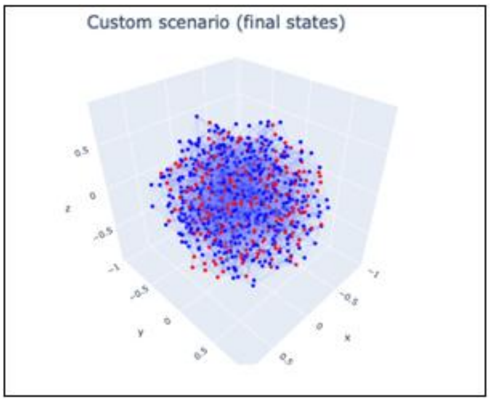


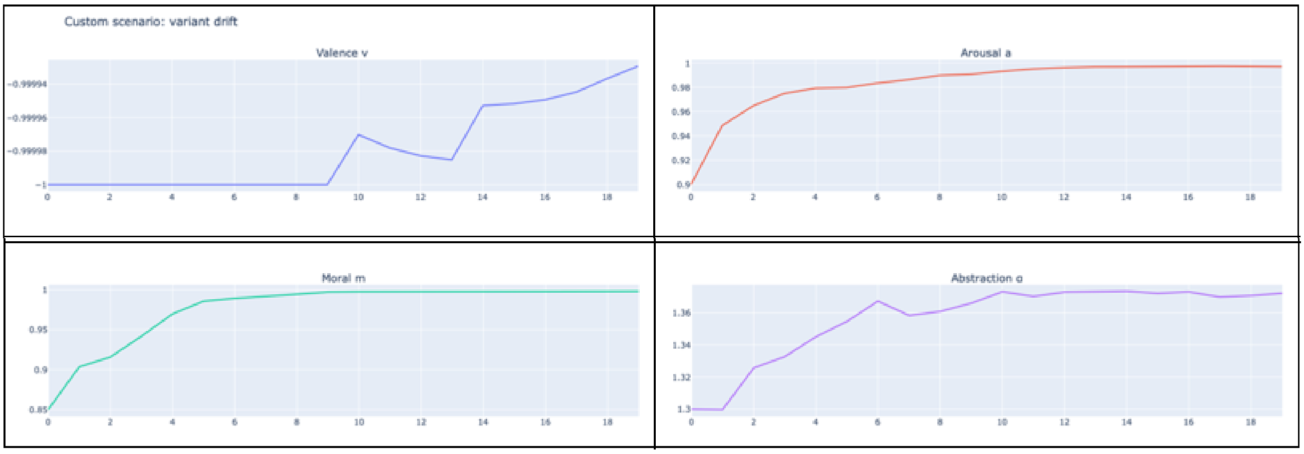


Appendix F

Utility: Build Your Own Experiment

Final informed fraction: 0.99 | R0≈ 88.00371485244976





Appendix G

Mathematical Proofs and Justifications for SCBA

This appendix supplies formal derivations, proof sketches, and identification arguments for all the mathematical components of the **Strategic Contagion with Behavioral Amplification (SCBA)** model. Symbols follow the main text unless redefined locally. Throughout, $\sigma(z)=\frac{1}{1+e^{-z}}$ denotes the logistic.

A. Micro Layer: Utility, Salience, and Transmission

A.1 Random-Utility Foundation for the Logistic Sharing Rule

Claim. The mapping $P_{ij}(T \mid \mathcal{I}) = \sigma\big(k\, E[U_i(T \mid \mathcal{I})]\big)$ arises from a random-utility choice with Gumbel noise (Quantal Response / Logit).

Setup. Agent i chooses $a \in \{T, Q\}$ (Transmit or Quiet). Deterministic utility difference:

$$\Delta u \equiv u(T) - u(Q) = E[U_i(T \mid \mathcal{I})].$$

Observed utilities add i.i.d. type-I extreme value noise $\varepsilon_T, \varepsilon_Q$ with scale $1/k$. Then

$$\Pr(T) = \Pr\{\Delta u + \varepsilon_T - \varepsilon_Q \geq 0\} = \sigma\big(k\,\Delta u\big).$$

(Proof: the difference of i.i.d. Gumbels is logistic; cf. standard logit/RUM derivation.)

□

Corollary A.1 (Monotonicity). P_{ij} is strictly increasing in $E[U_i(T \mid \mathcal{I})]$ and has constant marginal effect at the inflection point $E[U] = 0$ equal to $k/4$.

A.2 Multiplicative Salience Aggregator

Claim. If (i) status and instrumental benefits enter *multiplicatively* through independent channels of amplification (valence, arousal, abstraction) and (ii) log-benefits add, then

$$B'_{\text{Total}}(\mathcal{I}) = S(\mathcal{I})(B_{\text{St}} + B_I) + B_S, \qquad S(\mathcal{I}) = \lambda(v)\, A(\mathcal{I})\, \alpha(\mathcal{I}).$$

Derivation. Suppose

$$\log B'_{\text{amp}} = \theta_v(v) + \theta_a(a) + \theta_\alpha(\alpha) + \log(B_{\text{St}} + B_I).$$

Exponentiating gives $B'_{\text{amp}} = \exp\{\theta_v\}\exp\{\theta_a\}\exp\{\theta_\alpha\}(B_{\text{St}} + B_I)$. Set $\lambda = e^{\theta_v},\ A = e^{\theta_a},\ \alpha = e^{\theta_\alpha}$ to obtain the product form. Add B_S if social bonding is not subject to the same amplifiers. □

Remark. The product form is the most parsimonious continuous aggregator consistent with independence and log-linearity; alternatives reduce to it under a first-order log-linear approximation.

Appendix H

A.3 Negativity Multiplier from Prospect Theory

Claim. Under a prospect-theoretic value function

$$V(x)=\begin{cases}x^{\beta}, & x\geq 0,\\ -\lambda_L(-x)^{\beta}, & x<0,\end{cases}\quad \text{with } \lambda_L>1,\ \beta\in(0,1),$$

the *effective* marginal impact of negative information on the sharing benefit is $\lambda(v)=\lambda_N>\lambda_P\approx 1$ for positive information.

Justification. Let Δb be the perceived magnitude of the item's consequence (e.g., the norm/physical loss the gossip signals). The expected incremental value from communicating a loss $-|\Delta b|$ versus a gain $+|\Delta b|$ satisfies

$$\frac{|V(-|\Delta b|)|}{V(+|\Delta b|)}=\frac{\lambda_L(-|\Delta b|)^{\beta}}{(|\Delta b|)^{\beta}}=\lambda_L>1.$$

This ratio scales the benefit channel for negative items, yielding $\lambda_N=\lambda_L>1$ and $\lambda_P\approx 1.$ □

A.4 Arousal as Surprisal

Claim. If attention/status gains from sharing increase with *surprisal* $H(\mathcal{I})=-\log P(\mathcal{I})$, then a first-order admissible amplifier is

$$A(\mathcal{I})=1+\gamma\,H(\mathcal{I}),\quad \gamma>0.$$

Justification. Suppose expected incremental reach satisfies $\mathbb{E}[\text{reach}\mid\mathcal{I}]\propto e^{\theta H(\mathcal{I})}$ (exponential response to surprisal is maximum-entropy consistent for positive, scale-free sensitivity). A first-order Taylor expansion at $H=0$ gives $e^{\theta H}\approx 1+\theta H$. Relabel $\gamma=\theta.$ □

Remark. Using $A=e^{\gamma H}$ is equally compatible; we adopt the affine form for analytical transparency and empirical flexibility.

Appendix I

A.5 Linguistic Abstraction Factor $\alpha\geq 1$

Claim. Under the Linguistic Category Model, let categories $c\in\{\text{DAV},\text{IAV},\text{SV},\text{ADJ}\}$ (from concrete to abstract) have weights w_c with $1=w_{\text{DAV}}<w_{\text{IAV}}<w_{\text{SV}}<w_{\text{ADJ}}$. If a framing uses a mixture $p(c)$, define

$$\alpha\equiv\sum_c p(c)w_c\quad\Rightarrow\quad\alpha\geq 1,$$

with strict inequality whenever any mass lies above DAV.

Proof. Since all $w_c\geq 1$ and $p(c)$ is a simplex, $\alpha=\sum_c p(c)w_c\geq\sum_c p(c)\cdot 1=1$. Strict inequality holds if $p(c^{\star})>0$ for some $c^{\star}\neq\text{DAV}.$ □

Interpretation. Abstract frames imply enduring traits, lowering refutability and broadening applicability; modeling this as α amplifying $B_{\text{St}}+B_I$ is a reduced-form of that mechanism.

Paper ID: SR25823223650 DOI: https://dx.doi.org/10.21275/SR25823223650

A.6 Share-Threshold Inequality

Claim. With alternatives S (share), V (verify then share), D (debunk), Q (quiet), the *share region* is

$$\alpha_i \, \mathbb{E}[\text{reach}_i] + \sigma_i f(a,m) \;\geq\; \pi_i(1-p_i)q(\ell)F_i + c_S \quad \text{and} \quad U_i(S) \geq U_i(V), U_i(D).$$

Proof. Directly compare payoffs given in the main text:

$$U_i(S) = \alpha_i \mathbb{E}[\text{reach}_i] + \sigma_i f(a,m) - \pi_i(1-p_i)q(\ell)F_i - c_S,$$
$$U_i(Q) = 0.$$

Then $U_i(S) \geq 0$ yields the stated inequality. Enforce $U_i(S) \geq U_i(V), U_i(D)$ analogously by substituting their expressions. □

Comparative statics. Since $\partial U_i(S)/\partial a = \sigma_i f_a > 0$, $\partial U_i(S)/\partial m = \sigma_i f_m > 0$, $\partial U_i(S)/\partial \mathbb{E}[\text{reach}_i] = \alpha_i > 0$, the share region expands in arousal/moralization and centrality.

Appendix J

B. Meso Layer: Network Thresholds and Structure

B.1 Spectral Threshold for SIS-like Linearization

Claim. For a continuous-time SIS on graph with infection rate β and recovery δ, the disease-free equilibrium is unstable (allowing macroscopic spread) iff

$$\frac{\beta}{\delta}\,\lambda_1(A) > 1,$$

where $\lambda_1(A)$ is the spectral radius of the (nonnegative) adjacency matrix A.

Proof sketch. Let $x_i(t)$ be infection probability. The NIMFA linearization near $x = 0$ is $\dot{x} = (\beta A - \delta I)\,x$. Solutions satisfy $x(t) = \sum_k c_k e^{(\beta\lambda_k - \delta)t} v^{(k)}$. The leading mode grows iff $\beta\lambda_1 - \delta > 0$. Perron–Frobenius ensures λ_1 is real and associated with a nonnegative eigenvector. □

B.2 Independent Cascade (IC) and Bond Percolation Threshold

Claim. In IC with single-shot transmission probabilities T_{ij}, the cascade condition is governed by the spectral radius of the *transmission matrix* $T = (T_{ij})$; a sufficient condition for supercriticality is

$$\rho(T) > 1,$$

and in configuration-model networks with homogeneous per-edge transmission p, the branching approximation yields

$$\mathcal{R}_0 \approx p\,\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}.$$

Proof sketch. IC is equivalent to directed bond percolation with open-edge probability T_{ij}. Near criticality, the expected number of newly activated nodes reached by following an active half-edge equals the spectral radius of the non-backtracking operator; bounding by $\rho(T)$ suffices for a conservative threshold. For a configuration model with iid degrees, the Galton–Watson offspring mean after following a random edge is $p\,\mathbb{E}[K_{\text{excess}}] = p(\langle k^2 \rangle - \langle k \rangle)/\langle k \rangle$. Supercriticality requires this mean > 1. □

B.3 Bridges, Hubs, and Spectral Effects

Claim. Removing cross-community bridges weakly decreases $\lambda_1(A)$.

Proof. For nonnegative matrices, edge deletions correspond to entrywise decreases $A' \leq A$. By monotonicity of the spectral radius on the cone of nonnegative matrices (Collatz–Wielandt), $\lambda_1(A') \leq \lambda_1(A)$. $\square$

Corollary. Rate-limiting "weak ties" reduces $\mathcal{R}_0 = \frac{\beta}{\delta}\lambda_1(A)$ linearly in $\lambda_1(A)$.

Appendix K

C. Macro Implication: Valence Asymmetry in $\mathcal{R}_0$

C.1 From Salience to $\mathcal{R}_0$

Claim. If negative items $\mathcal{I}_N$ have $S_N > S_P$ (via $\lambda_N > \lambda_P,\ A_N \geq A_P,\ \alpha_N \geq \alpha_P$) and the micro mapping $E[U] \mapsto P_{ij}$ is strictly increasing, then (holding network fixed)

$$\beta(\mathcal{I}_N) > \beta(\mathcal{I}_P) \quad \text{and typically} \quad \delta(\mathcal{I}_N) \leq \delta(\mathcal{I}_P),$$

hence

$$\mathcal{R}_0(\mathcal{I}_N) = \frac{\beta(\mathcal{I}_N)}{\delta(\mathcal{I}_N)}\lambda_1(A) \ > \ \mathcal{R}_0(\mathcal{I}_P).$$

Proof. By A.1–A.2, S increases $E[U]$, which increases P_{ij}; aggregating over edges yields higher effective β. High-arousal content sustains engagement (argument persistence), so δ does not increase and may decrease. The ratio thus rises. $\square$

Critical-zone corollary. There exists parameter ranges where $\mathcal{R}_0(\mathcal{I}_P) \leq 1 < \mathcal{R}_0(\mathcal{I}_N)$, producing the observed asymmetry in cascade realization.

Appendix L

D. Evolutionary Extension: Replicator–Mutator and Error Threshold

D.1 Replicator–Mutator Fixed Point

Let x index message variants with fitness $w(x) > 0$ and mutation kernel $Q(y \mid x)$ (row-stochastic). The population update is

$$p_{t+1}(y) = \frac{1}{\bar{w}_t}\sum_x p_t(x)\, w(x)\, Q(y \mid x), \qquad \bar{w}_t = \sum_x p_t(x) w(x).$$

Claim. A stationary distribution p^* exists and is the Perron–Frobenius (PF) right eigenvector of the positive matrix $M = WQ$ (with $W = \mathrm{diag}(w)$), normalized to sum to 1.

Proof. M is primitive if Q is primitive and $w(x) > 0$ for all x. By PF, M has a unique positive eigenvector v for the spectral radius $\rho(M)$. Writing the update in vector form: $p_{t+1} = \frac{M^\top p_t}{\mathbf{1}^\top W p_t}$. At stationarity, $p^* \propto v$. Uniqueness follows from primitivity. $\square$

D.2 Error Threshold (Two-Type Illustration)

Consider types $\{x_0, x_1\}$ with $w_0 > w_1 > 0$ and symmetric mutation $Q = \begin{bmatrix} 1-\mu & \mu \\ \mu & 1-\mu \end{bmatrix}$.

Claim. The stationary mass on the fittest "master" type x_0 is

$$p_0^* = \frac{w_0(1-\mu) - w_1\mu}{w_0 + w_1 - 2\mu(w_0 + w_1)}.$$

There exists $\mu_c \in (0, 1/2)$ such that $p_0^* \downarrow 1/2$ and then declines rapidly as $\mu \uparrow \mu_c$ ("error threshold").

Proof. Compute $M = WQ$, then the (normalized) PF right eigenvector; the expression above follows by algebra. Solve $p_0^* = 1/2$ for μ to obtain $\mu_c = \frac{w_0 - w_1}{2(w_0 + w_1)}$. For $\mu > \mu_c$, selection cannot maintain concentration on x_0. □

Interpretation. In SCBA, μ increases with memory pressure and pragmatic exaggeration; crossing μ_c collapses semantic integrity.

Appendix M

D.3 Rate–Distortion Form for Memory Kernel

Claim. The memory/compression transition $Q_{\text{mem}}(y \mid x) \propto p(y)\, e^{-\beta d(x,y)}$ solves a constrained optimization: among all channels $Q(\cdot \mid x)$ with bounded expected distortion, choose the one maximizing expected (log) prior probability—equivalently, minimize expected distortion plus a KL regularizer.

Derivation. Minimize

$$\mathcal{L} = \sum_x p(x) \sum_y Q(y \mid x) \Big[d(x,y) + \frac{1}{\beta} \log \frac{Q(y \mid x)}{p(y)} \Big]$$

over row-stochastic Q. The Euler–Lagrange/KKT condition yields $Q(y \mid x) \propto p(y)\, e^{-\beta d(x,y)}$, normalized per x. This is the Gibbs form of the rate–distortion solution (and of the Information Bottleneck with distortion d). □

D.4 RSA-Style Pragmatic Kernel

Claim. If speakers choose utterances u to maximize a utility

$$U(u; s) = \alpha \log L_0(s \mid u) - c(u) + \zeta\, \text{social}(u),$$

and sample with Luce/soft-max, then

$$Q_{\text{prag}}(y \mid x) \;\propto\; \exp\{\lambda\, U(u_y; s_x)\},$$

inducing upward drift in arousal a and moral-emotional intensity m when these raise U.

Proof. The Luce/soft-max rule gives $\Pr(u \mid s) \propto e^{\lambda U(u;s)}$. Mapping utterances u to feature changes $x \to y$ produces the stated kernel. If $\partial U / \partial a > 0, \partial U / \partial m > 0$, then the transition has positive drift in those coordinates. □

Appendix N

E. Identification and Estimation (Sketch)

E.1 Mapping Language to $\beta(\ell)$ and $\delta(\ell)$

Claim. With observed exposure graph and time-stamped re-shares, $\beta(\ell)$ and $\delta(\ell)$ are nonparametrically identified up to scale under standard hazard assumptions.

Sketch. Estimate a discrete-time hazard for first re-share by neighbor j conditional on exposure by i:

$$\Pr(\text{share}_{ij,t} = 1 \mid \text{not shared yet}, \ \ell) = \sigma(\gamma_0 + \gamma^\top \ell + \text{controls}).$$

This recovers $\beta(\ell)$ up to exposure scaling. A separate stopping hazard on active spreaders recovers $\delta(\ell)$ Normalization fixes the scale (e.g., set average w_{ij} to 1). □

E.2 Recovering the Salience Channels

Claim. If $\ell = (v, a, m, n, u, \text{abstract})$ enters $\ln \beta(\ell) = \beta_0 + \beta_a a + \beta_m m + \beta_n n + \beta_u u + \beta_v \mathbf{1}\{v < 0\} + \beta_\alpha \,\text{abstract}$, then the *partial* effects identify the components of $S = \lambda A \alpha$ under mean-independence of regressors.

Sketch. The linear index yields separable marginal contributions; each coefficient provides a local elasticity. Interactions can be added to test multiplicativity (e.g., significance of $v \times a$, $v \times \text{abstract}$). □

Appendix O

F. Additional Comparative Statics and Robustness

F.1 Supermodularity and Strategic Complements

Claim. If the sharing payoff $U_i(S)$ is supermodular in $(a, m, \mathbb{E}[\text{reach}_i])$, increases in any one variable raise the marginal return to the others, steepening diffusion gradients.

Proof sketch. Let $U_i(S) = g(a, m) + h(\mathbb{E}[\text{reach}_i]) - \text{penalty}$, with $g_{am} > 0$, $h'' \geq 0$. Then $\partial^2 U / \partial a \, \partial m > 0$ and $\partial^2 U / \partial a \, \partial \mathbb{E}[\text{reach}_i] > 0$ (via composition with a convex h). Topkis's theorem implies monotone comparative statics in parameters that shift these arguments. □

F.2 Invariance and Normalizations

- Scaling k and $E[U]$ by reciprocal constants leaves P_{ij} unchanged: $\sigma(kE[U]) = \sigma(\tilde{k}\,\tilde{U})$ with $\tilde{k} = ck, \tilde{U} = E[U]/c$.
- Multiplying all benefits and costs by a common factor rescales k, not behavior, under logit normalization.

Appendix P

G. From Micro to Macro: Putting It Together

Theorem (SCBA cascade dominance). Suppose:

1. $S_N > S_P$ in expectation for negative vs. positive items (A.2–A.5);
2. $E[U] \mapsto P_{ij}$ is strictly increasing (A.1);
3. $\beta(\ell)$ increases and $\delta(\ell)$ weakly decreases in $(a, m, n, u, \mathbf{1}\{v < 0\}, \text{abstract})$;
4. The network satisfies $\lambda_1(A) > 0$.

Then there exists a parameter region where

$$\mathcal{R}_0(\mathcal{I}_P) \leq 1 < \mathcal{R}_0(\mathcal{I}_N),$$

so negative content produces macroscopic cascades while positive content does not.

Proof. By (1)–(2), $P_{ij}^N > P_{ij}^P \Rightarrow \beta_N > \beta_P$. By (3), $\delta_N \leq \delta_P$. Fix A; continuity in parameters implies there is a region where $\frac{\beta_P}{\delta_P}\lambda_1 \leq 1 < \frac{\beta_N}{\delta_N}\lambda_1$. $\square$