Social Physics: A Formal Framework for Quantifying Perceptual Distortion and Predicting Social Fission

Sara Skouri

November 2025

Abstract

This paper establishes **Social Physics** as a formal, predictive science of digital perception. We introduce the **Perceived Exaggerated Amplification (P-E-A)** framework—a composite metric quantifying distortion between algorithmic visibility and physical-world baselines. By defining the **Critical Distortion Threshold (CDT)** and cascade reproduction metric R_{fission}, we model societal breakdowns as measurable phase transitions. The work presents four amplification components (prevalence, emotional intensity, coordination, and bias skew), formal dynamical laws, and a research program to defend **perceptual integrity**. Social Physics bridges computational social science, ethics, and dynamical systems theory, offering tools to predict and prevent runaway social cascades in digital ecosystems.

Contents

1 The Spark: A Portrait of Dissonance

This work began with a persistent observation: the digital world presented in social feeds appeared systematically louder, angrier, and more unanimous than offline reality. Scrolling through platforms revealed a landscape where nuanced discourse vanished beneath floods of emotionally charged content. Every post seemed amplified—not in proportion to real-world prevalence, but in proportion to the engagement algorithms governing them. This created a quiet but constant cognitive dissonance: the digital environment felt skewed, harsher, and less forgiving than the complex, moderate reality experienced offline.

That disconnect sparked the central research question: *If the digital environment feels so different from lived reality, is it revealing a true societal shift—or a distorted mirror of perception?* And if it is a distortion, what measurable mechanisms produce it? This paper traces the journey from that initial dissonance to a formal, predictive science: Social Physics.

2 The Diagnosis: Perceived Exaggerated Amplification (P-E-A)

To diagnose the dissonance between online perception and offline reality, this work introduces the concept of **Perceived Exaggerated Amplification (P-E-A)**. P-E-A describes the systematic distortion in which algorithmically curated environments inflate the perceived prevalence and intensity of certain viewpoints, creating the illusion of consensus around extreme or emotionally charged content. In other words, it is the hidden "ghost" shaping our perception of social reality.

2.1 The Three Mechanisms of P-E-A

This phenomenon arises from three interconnected mechanisms:

1. **Algorithmic Incentive:** Platform architectures are designed to maximize engagement. Content that provokes strong emotional reactions—outrage, fear, moral indignation—receives amplified visibility through ranking and recommendation systems, artificially boosting its apparent significance.

2. **Psychological Vulnerability:** Human cognition is naturally drawn to salient, high-arousal information. Biases like confirmation bias and attentional preference make users more likely to notice, engage with, and remember extreme content, mistaking its visibility for widespread acceptance.

3. **The Perceptual Loop:** Repeated exposure to amplified content shapes users' beliefs about what is normal or widely endorsed. This perceived consensus then influences their

own engagement—liking, sharing, commenting—creating a feedback loop that signals the algorithm to further promote such content, reinforcing the distortion.

Together, these mechanisms form the **P-E-A feedback loop**, a self-perpetuating cycle of digital perception distortion. It is this engineered cycle—more than individual bias or preference—that constitutes the hidden architecture behind modern perceptual confusion.

2.2 Formal Components of Digital Distortion

Building on these mechanisms, we formalize P-E-A into four measurable, orthogonal dimensions:

2.2.1 Prevalence Amplification (A_p)

The ratio of a topic's platform visibility to its real-world baseline prevalence:

$$A_p(X,t) = \frac{\hat{p}_{\text{feed}}(X,t)}{\hat{p}_{\text{base}}(X)}$$

2.2.2 Emotional Intensity Amplification (A_e)

The ratio of average emotional arousal in feeds versus a neutral baseline:

$$A_e(X,t) = \frac{E_{\text{feed}}(X,t)}{E_{\text{base}}(X)}$$

2.2.3 Bot & Coordination Amplification (A_b)

Quantifies inauthentic or highly coordinated activity:

$$A_b(X,t) = 1 + \gamma \cdot \left(\frac{B_{\text{feed}}(X,t)}{B_{\text{base}}} \right)$$

2.2.4 Bias Skew Amplification (A_s)

Measures systematic slant and diversity loss in amplified sources:

$$A_s(X,t) = 1 + \delta_1 \cdot |\mu_{\text{feed}} - \mu_{\text{base}}| + \delta_2 \cdot \max(0, \sigma_{\text{base}} - \sigma_{\text{feed}})$$

2.3 Composite P-E-A Score

The overall P-E-A score is computed as the root-mean-square of its four components, ensuring scale invariance and sensitivity to extremes:

$$\text{P-E-A}(X,t) = \sqrt{\frac{A_p(X,t)^2 + A_e(X,t)^2 + A_b(X,t)^2 + A_s(X,t)^2}{4}}$$

This formulation captures the combined magnitude of perceptual distortion while avoiding arbitrary weighting of components, transforming the concept from theoretical model to auditable, technical reality.

3   The Evidence: The Consilience of Distortion

The power of the P-E-A framework lies in its ability to integrate evidence across disparate fields into a unified explanation for digital-age phenomena. This consilience reveals P-E-A not as a niche observation, but as a fundamental force reshaping modern society.

3.1   The Political Shock: Quantifying Democratic Distortion

The 2016 Brexit referendum serves as a primary case study. Pre-vote polls consistently showed a narrow lead for "Remain," yet the "Leave" campaign secured a victory that defied expectations. A P-E-A analysis reveals a critical distortion: while online discourse was saturated with highly visible, emotionally charged pro-Leave content, this amplified visibility was not proportional to underlying support. By quantifying the gap between algorithmic visibility and actual voter intention, the P-E-A framework provides a measurable explanation for the polling failure and resultant political shock, demonstrating how perception can be systematically manipulated.

3.2   The Psychological Toll: The Mechanisms of Digital Anxiety

P-E-A translates from a systemic phenomenon to an individual cognitive burden. Constant exposure to an amplified reality of conflict and extremity activates innate cognitive biases, such as negativity bias and the availability heuristic, leading individuals to overestimate social threat and hostility. This creates a state of chronic vigilance and erodes perceived self-efficacy. P-E-A thus links platform design directly to measurable declines in mental well-being.

3.3   The Technological Core: From Diagnosis to Solution

Computer science provides the tools to operationalize P-E-A. The proposed **P-E-A Score** is a composite metric derived from platform data, quantifying distortion through sub-scores for prevalence amplification, emotional intensity, and network coordination. This enables inter-

ventions such as "diversity-of-opinion boosts" in ranking algorithms. Here, ethical reflection meets actionable design.

3.4   The Philosophical Crisis: The Right to Perceptual Integrity

Ultimately, P-E-A demands a philosophical reckoning. By systematically distorting the informational environment, platforms undermine **perceptual integrity**—the right of individuals to form beliefs in an environment free from engineered manipulation. This constructs a modern analogue to Plato's Cave, where the "shadows" are algorithmically cast. The ethical imperative shifts from merely defending free speech to actively defending the epistemic commons on which democratic deliberation depends.

3.5   The Media Evolution: From Lippmann's Pictures to Algorithmic Feeds

The distortion of perception is not a purely digital phenomenon—it has deep historical roots. Walter Lippmann described the "pictures in our heads," the simplified, often misleading mental representations of public affairs shaped by media. P-E-A extends this insight to the digital age: algorithmic feeds, social amplification, and personalized ranking systems are the modern instruments creating these pictures on a massive, real-time scale. By tracing the evolution from traditional media to algorithmically curated content, P-E-A demonstrates that the mechanisms of perceptual distortion are not incidental, but systemic and cumulative, affecting the human experience across decades and platforms.

Together, these lenses establish P-E-A as a robust and necessary framework. It unites electoral surprises, psychological distress, algorithmic design, and foundational ethics, proving that the distortion of perception is the central challenge of the digital public square.

4   The Timing Act: How We Arrived at Social Physics

Every scientific field begins as a problem of language—a struggle to name a phenomenon that humanity can feel but not yet measure. Sociology emerged when thinkers like Durkheim and Weber sought the hidden patterns behind social order and conflict. Psychology formalized the study of the mind that drives that collective behavior. For centuries, these disciplines provided the essential tools to understand a world bounded by physical interaction.

Then came the inflection point: the information age. In this new era, computers became social actors, data became culture, and human perception itself became a programmable substrate. The traditional tools of the social sciences, powerful as they were, began to fracture under the strain. They could describe the symptoms—rising polarization, societal anxiety, viral misinformation—but they could not predict its trajectory or quantify its underlying

mechanics. The very architecture of reality had changed, and a new science was needed to map it.

The development of the Perceived Exaggerated Amplification (P-E-A) framework marked the critical turning point. For the first time, the distortion of social perception could be expressed not just as a concept, but as a measurable, composite quantity.

Social Physics is the completion of that arc. It represents the natural convergence of psychology, sociology, and the principles of physics into a unified science of digital perception and collective behavior. It is not a rejection of the disciplines that came before, but their necessary evolution.

Every equation, every threshold, every law in this document is part of a single, long story: the story of humanity learning to see its own reflection in the mirror of its own creations, and finally developing the language to understand what it sees.

5 The Physics of Perception: A Nuclear Analogy

This framework is inspired by the conceptual power of nuclear physics to reveal the hidden rules governing the stability and collapse of matter. Social Physics seeks to do the same for the stability and collapse of societal consensus.

5.1 Parallel Structures

- **Systems of Bound Energy:** Both atoms and human networks are systems of immense latent energy, bound by fragile forces of cohesion. In societies, this energy is cognitive and emotional, bound by shared identity, trust, and norms.

- **The Quest for Stability:** An atom seeks a stable state in its nucleus. A person, and by extension a community, seeks a stable state in their perceptual environment—a reliable sense of what is real and true.

- **Critical Thresholds:** In nuclear physics, a critical mass is the minimum amount of fissile material needed to sustain a chain reaction. In Social Physics, the Critical Distortion Threshold (CDT) is the minimum level of perceptual distortion required to trigger a runaway social cascade, or Social Fission.

- **The Fission Process:** When an atom's nucleus is overloaded with external neutrons, it splits, releasing tremendous energy. When a social system is overloaded with perceptual distortion (P-E-A), it can also split—not into nuclei, but into conflicting, hostile realities, releasing the energy of societal conflict.

5.2 Comparative Ontology

Domain	Nuclear Physics	Social Physics
Fundamental Unit	Atom	Individual Agent (φ_i)
Energy Source	Nuclear binding energy	Emotional-perceptual energy
Trigger	Neutron impact	Exposure to high-P-E-A content
Critical Threshold	Critical mass	Critical Distortion Threshold (CDT)
Cascade Metric	Neutron reproduction rate (R)	R_{fission}
Containment	Control rods	Algorithmic dampening, diversity circuits

Table 1: Comparative ontology between nuclear physics and social physics

5.3 From Metaphor to Mathematics

While the nuclear analogy provides conceptual scaffolding, Social Physics operationalizes these parallels into measurable constructs:

$$R_{\text{fission}} = \rho(\mathbf{D} + \mathbf{W}), \quad \text{CDT} = \inf\{\alpha \geq 0 : \rho(\mathbf{D} + \mathbf{W}(\alpha)) \geq 1\}$$

where $\mathbf{D}$ is the decay matrix and $\mathbf{W}$ the influence matrix parameterized by P-E-A.

6 Why Social Science and Physics Make a Powerful Fusion

This integration is not a direct one-to-one mapping, but a powerful conceptual scaffold. It allows us to transpose a mature science of thresholds, chain reactions, and energy release onto the complex dynamics of digital society.

Social systems and physical systems may seem worlds apart—one built from consciousness and culture, the other from particles and forces—but they share a deep, structural architecture of interaction, energy exchange, and emergent order.

Physics provides a language for dynamics: it models how and why systems evolve, stabilize, or collapse. Social science provides a language for meaning: it describes the beliefs, identities, and norms that constitute human experience. Alone, each is powerful but incomplete for the digital age. Sociology can describe the pattern of polarization; psychology can name the resulting anxiety. But without the dynamic laws of physics, they cannot predict the tipping point.

Their fusion is powerful because it is complementary. Social Physics does not seek to reduce humanity to mere equations, but to use the precision of physics to reveal the universal logic underlying our collective behavior. By adopting the rigorous concepts of energy transfer, diffusion, and critical thresholds, we can quantify the invisible pressures that build within a society. We can move from observing that a system is "under stress" to calculating the exact moment when perceptual distortion crosses the Critical Distortion Threshold, triggering the runaway cascade of Social Fission.

This fusion is not just intellectually elegant; it is urgently necessary. It brings testability and predictive power to the most profound human questions: How do perceptions crystallize into reality? When does disagreement become dangerous? Where sociology maps the landscape of meaning and psychology studies the individual mind, Social Physics provides the mathematical skeleton—the hidden symmetry that binds chaos and coherence, allowing us to finally see the architecture of our own collective existence.

7 The Leap: The Formal Foundations of Social Physics

The cross-disciplinary evidence demonstrates that Perceived Exaggerated Amplification (P-E-A) is a pervasive and consequential feature of digital society. This part synthesizes these insights into a formal, predictive science: Social Physics. This new field moves beyond description to model the fundamental laws governing how information flows, perceptions form, and societies reach tipping points.

7.1 Core Mission Statement

The mission of Social Physics is to create measurable laws of digital perception to predict and prevent runaway social cascades.

7.2 The New Lexicon

To enable precise discourse, we define the field's core constructs:

- **Critical Distortion Threshold (CDT):** The minimum level of P-E-A for a given topic within a specific network at which a cascade becomes self-sustaining.

- R_{fission} **(Cascade Reproduction Number):** The average number of new hostile or coordinated mobilization events spawned by a single prior event. It measures the contagiousness of a cascading social reaction.

- **Social Fission:** The catastrophic, runaway process that occurs when $R_{\text{fission}} > 1$, whereby a shared reality splits into conflicting perceptual realities, releasing energy in the form of societal conflict and mobilization.

9

7.3   The Foundational Laws & Equations

Social Physics is grounded in formal, testable propositions that define the relationships between distortion, cognitive energy, and societal state change.

7.3.1   The First Law of Social Physics (The Law of Amplified Perception)

For a given topic in a digitally-mediated environment, the Perceived Exaggerated Amplification (P-E-A) is a function of the product of its algorithmic engagement weight and its innate emotional valence, and is inversely proportional to the perceptual diversity of the information environment.

This law establishes that distortion is not random but is systematically driven by platform architecture and human psychology.

7.3.2   The Second Law of Social Physics (The Law of Critical Thresholds)

Every sociotechnical system has a Critical Distortion Threshold (CDT) for a given topic. A state of Social Fission is reached when the P-E-A score meets or exceeds the CDT, making a large-scale cascade statistically inevitable.

This law posits that societal systems have measurable breaking points, analogous to phase transitions in physical systems.

7.3.3   The Operational Equations

1. **The P-E-A Score:** The fundamental metric of perceptual distortion, calculated as the root mean square of its core components to avoid arbitrary weighting and capture their combined magnitude:

$$\text{P-E-A}(X,t) = \sqrt{\frac{A_p(X,t)^2 + A_e(X,t)^2 + A_b(X,t)^2 + A_s(X,t)^2}{4}}$$

2. **The R_{fission} Equation (Cascade Metric):** The key variable for predicting cascade dynamics:

$$R_{\text{fission}}(X,t) = \mathbb{E}\left[\frac{N_{\text{events}}(X,t+\Delta t)}{N_{\text{events}}(X,t)} \,\middle|\, \text{P-E-A}(X,t), \Theta\right]$$

   Interpretation:

   - $R_{\text{fission}} < 1$: The cascade is subcritical and dying out.
   - $R_{\text{fission}} = 1$: The cascade is at critical threshold (CDT).
   - $R_{\text{fission}} > 1$: The cascade is supercritical (Social Fission).

3. **The Critical Distortion Threshold (CDT) Equation:** The formal definition of the system's tipping point:

$$\boxed{\mathrm{CDT}(X) = \inf\{\alpha \geq 0 : R_{\mathrm{fission}}(X \mid \text{P-E-A} = \alpha) \geq 1\}}$$

In words: The CDT is the smallest P-E-A score for which the topic X becomes self-sustaining and enters a runaway cascade.

With these foundations, Social Physics provides the formal framework to transition from observing distortion to quantitatively modeling its trajectory and predicting its most dangerous consequences.

8 The Fortress: Stress-Testing the Paradigm

A robust scientific framework must not only explain phenomena but also withstand rigorous scrutiny. To build credibility and refine the paradigm, we here preemptively engage with the most substantial criticisms that Social Physics and the P-E-A framework are likely to confront.

1. **"This is just repackaged 'echo chamber' or 'polarization' theory with fancier math."**

   Our Defense: This critique confuses outcomes with processes. Echo chambers describe the separation of information environments, and polarization describes the ideological distance between groups. P-E-A explains the internal dynamic that occurs within these states: the systematic **inflation of perceived consensus** that makes a minority view feel like a dominant majority. The mathematics is not mere decoration; it is essential for quantifying the distortion's intensity and predicting its critical threshold, moving from qualitative description to quantitative prediction.

2. **"The P-E-A Score and CDT are arbitrary composites. You're cherry-picking metrics."**

   Our Defense: The components of the P-E-A Score are not arbitrary; they are derived from foundational constructs in their respective fields: prevalence from communication theory, emotional intensity from psychology, and coordination from network science. We combine them using a Root Mean Square (RMS) specifically to **avoid arbitrary weighting** and to capture their combined magnitude. The methodology for estimating the CDT—through breakpoint regression analysis of real-world hostility indicators against the P-E-A score—is a standard statistical technique for identifying thresholds in complex systems.

3. "You're anthropomorphizing algorithms. They just reflect user preference."

Our Defense: This misattributes the core argument. We do not claim algorithms have intent, but that their design creates **systemic, emergent effects**. Algorithms absolutely reflect user engagement. However, the *outcome* is a distortion of *perceived prevalence.* A small, highly engaged group can generate the visibility of a massive majority. The "distortion" is rigorously defined as the measurable gap between the *volume of a signal* and the *size of the group it represents* in reality.

4. "This is too reductive. Human society can't be modeled like a physical system."

Our Defense: Social Physics is not reductionist but **abstractive**. It does not seek to explain the rich, complex psychology of an individual—why one person becomes a "fool" while another becomes a sage. That is the vital domain of individual psychology. Instead, Social Physics operates at the **collective level**, asking a different question: Under what systemic conditions do *populations* become predictably more prone to cascading behaviors? It identifies the **scaffolding of collective behavior**—the system-level conditions under which beliefs become unstable at scale. Just as epidemiology doesn't model the poetry of a single life but can predict the spread of a virus through a population, Social Physics seeks the laws governing the "epidemiology of perception."

5. "Interdisciplinary vagueness dilutes rigor."

Our Defense: On the contrary, cross-disciplinary synthesis is what makes rigor portable. We do not blend fields vaguely; we **translate methods across them with explicit assumptions**. Each component of the P-E-A score maintains its disciplinary integrity while contributing to a composite metric, keeping conceptual boundaries visible and testable. This framework provides a precise, shared language for psychologists, sociologists, and computer scientists to collaborate on a single problem, enhancing rigor rather than diluting it.

6. "It's just metaphorical physics."

Our Defense: Metaphor is the starting line for insight, not the finish line for science. While the nuclear analogy provides a powerful conceptual scaffold, concepts like the **Critical Distortion Threshold (CDT)** and the cascade metric $R_{\mathbf{fission}}$ are fully **operationalized into measurable constructs**. They are anchored in mathematical formulation and statistical testing, transforming a suggestive metaphor into a falsifiable, predictive model.

7. "The 'Early-Warning System' is a slippery slope to censorship."

Our Defense: This is a vital ethical concern. The core design principle of our proposed system is that it diagnoses the **environment, not the content**. The output is not "remove this post," but an alert that "the perceptual environment for Topic X is nearing a critical threshold for cascade risk." The prescribed response is not censorship, but **corrective transparency**—such as boosting viewpoint diversity, issuing context labels, and promoting media literacy. The system is a canary in the coal mine, designed to support informed democratic deliberation, not to supplant it.

9 The Blueprint: A Research Program for Reality Integrity

With the conceptual framework established, we now present a concrete research program to transition Social Physics from theory to practice. This blueprint outlines foundational studies, operational tools, and ethical guardrails necessary to build a rigorous science of perceptual integrity.

9.1 Foundational Studies: The First Five Pillars

Study 1: Validation of the CDT

> *"Quantifying the Critical Distortion Threshold: A Forensic Analysis of the Brexit Referendum."*
> **Goal:** Empirically estimate the CDT using historical data, establishing the predictive validity of the P-E-A score by analyzing correlations with real-world mobilization, opinion shifts, and polling discrepancies.

Study 2: The Cognitive Cost of Amplification

> *"The Amplified Mind: A Longitudinal Study of P-E-A Exposure and Psychological Outcomes."*
> **Goal:** Causally link exposure to high-P-E-A environments with measurable increases in state anxiety, social cynicism, and reduced self-efficacy in a controlled setting.

Study 3: Engineering Perceptual Integrity

> *"Operationalizing the Perceptual Integrity Index: A/B Testing Platform Interventions."*
> **Goal:** Develop and test algorithmic adjustments (e.g., diversity-of-opinion boosts, temporal smoothing) designed to reduce the P-E-A score without compromising user engagement.

Study 4: Predicting Cascades in Real-Time

"From P-E-A to Social Fission: Building an Early-Warning Dashboard for Civil Society."

Goal: Create a prototype dashboard that monitors real-time P-E-A scores for high-risk topics and triggers alerts as metrics approach the estimated CDT.

Study 5: Formalizing the Field

"Social Physics: Foundational Axioms and a Research Agenda."

Goal: Publish the seminal theoretical paper defining the field's principles, scope, and challenges, establishing a shared language for interdisciplinary research.

9.2   Tangible Tools for a New Reality

- **The P-E-A Score:** An open-source, publicly documented metric for auditing platform health.

- **The Perceptual Integrity Index (PII):** A standardized, cross-platform scorecard for regulators and policymakers, evaluating the degree to which platforms preserve undistorted information environments.

- **The Social Fission Early-Warning System:** A dashboard providing journalists, civil society organizations, and platform safety teams with a data-driven lead indicator of emerging societal risks.

9.3   Embedded Ethical Guardrails

- **Privacy by Design:** All data aggregated and anonymized; no individual-level tracking or profiling.

- **Human-in-the-Loop:** Algorithmic flags serve human review and context-aware intervention, never automated content removal.

- **Transparency and Auditability:** Methodologies are fully open for public scrutiny and independent verification.

- **Dual-Use Mitigation:** Research outputs are designed defensively to protect public discourse, with clear documentation of potential misuse risks.

A field that seeks to measure and influence social perception must embed ethical design into its very methods. This blueprint charts a course from foundational research to real-world impact. Its goal is not merely to understand society, but to equip researchers, policymakers, and civil society with the tools to safeguard it.

10   The Invitation: On the Shoulders of Giants

This document began with a question born of personal dissonance. It now concludes with a vision forged in collective purpose. Social Physics does not arrive as a finished edifice, but as an open construction site—a foundation upon which we can build a deeper understanding of ourselves in this new digital epoch.

The journey from a single observation to a formal science has been one of synthesis, standing on the shoulders of giants across philosophy, psychology, political science, and physics. This work is a testament to the power of collaborative curiosity and is itself the product of profound guidance. It was the incisive mind and generous mentorship of **Mr. Carey Glenn Butler** that helped fan the initial spark of this synthesis into a steady flame, providing the mathematical rigor and visionary belief that propelled it forward.

In the same spirit of open collaboration, we extend this work as an invitation:

- We invite the **empiricist** to test these equations.

- We invite the **engineer** to build these tools.

- We invite the **ethicist** to strengthen these guardrails.

- We invite the **psychologist and sociologist** to deepen the human context within this mathematical scaffold.

The central challenge of our time is not a lack of information, but a crisis of perception. The mission of Social Physics—to create measurable laws of digital perception to predict and prevent runaway social cascades—is therefore a profoundly human one. It is a commitment to preserving our shared reality.

This is not the end. It is the very first step. The blueprint is drawn. The foundations are laid.

Now, let us build.

Made with red♡ by Sara Skouri

References

[1] Durkheim, É. (1893). *The division of labour in society.* Macmillan.

[2] Lippmann, W. (1922). *Public opinion.* Harcourt, Brace and Company.