

Rethinking Public Trust: A Preliminary Field-Governance Perspective

Wanhong HUANG
huangwanhong@serendip.ngo

Abstract

Public trust is often approached as an attitude held by individuals, a property of institutions, or a resource to be increased through transparency, credibility, and repeated cooperation. This paper reconsiders public trust as a historically evolving relational condition whose formation, stabilization, and revision depend upon interacting social, institutional, interpretive, symbolic, and embodied processes. Trust and distrust are treated as potentially self-reinforcing configurations: once established, they can alter subsequent interaction, shape the interpretation of evidence, redistribute interpretive authority, and contribute to the production of further conditions that support their continuation. This dynamic makes both trust and distrust normatively ambivalent. Trust can sustain cooperation while also supporting dependency, collusion, extraction, or interpretive closure; distrust can preserve critical distance while also becoming self-sealing and resistant to corrective experience. The paper introduces a preliminary field-governance perspective as a limited analytical device for examining these coupled dynamics. Rather than proposing a general theory of field governance, it considers how a designed governance field may perturb already existing relational conditions and interact with trust, distrust, evidence, status, attention, and interpretive power through higher-order effects. On this basis, public balance is approached neither as equalized influence nor as maximal trust, but as a revisable and scale-sensitive condition in which no stabilized configuration permanently forecloses alternative evidence, interpretation, participation, or relational exit. The paper concludes by developing public re-cognizability as a provisional governance concern: public arrangements should preserve the capacity to reconstruct how judgments formed, reopen settled interpretations, and revise trust and distrust as relational histories change.

Keywords: public trust; distrust; field governance; relational dynamics; attractor-like stabilization; interpretive power; public epistemology; re-cognizability; generative relational theory

Discussion Paper Note

This paper is a preliminary conceptual inquiry into public trust from a field-governance perspective. Its object is deliberately narrower than a general theory of field governance. The discussion uses public trust and distrust as a concrete domain in which historically accumulated relations, interpretive authority, evidence, status, institutional arrangements, symbolic conditions, and governance interventions can become mutually coupled.

The paper begins from the observation that trust and distrust cannot be reduced to isolated psychological attitudes. They can emerge from repeated interaction, shared experience, institutional guarantees, security mechanisms, public reputation, symbolic authority, sensory and affective conditions, and other features of a relational environment. Once formed, they can also modify future relations. Trust may lower defensive barriers, enable cooperation, change the interpretation of ambiguous evidence, and generate further experiences that stabilize trust. Distrust can similarly alter interaction, reduce opportunities for corrective experience, and produce interpretations and responses that reinforce distrust. The paper therefore treats both as historically evolving and potentially attractor-like relational configurations.

Attractor language is used cautiously. The paper does not assume that social relations instantiate a fixed dynamical system or that trust corresponds to a single mathematical attractor. The purpose of the language is to clarify path dependence, feedback, hysteresis, stabilization, relational lock-in, and the possibility that small differences in initial conditions can be amplified by subsequent interaction. The relevant dynamics may themselves evolve as relations, institutions, and interpretive practices change.

The normative analysis likewise avoids identifying trust with social good or distrust with social failure. Trust can support cooperation, care, coordination, and institutional continuity, while also enabling dependency, collusion, extraction, exclusion, or the consolidation of interpretive privilege. Distrust can preserve critical distance, verification, and resistance to concentrated power, while also becoming self-sealing and resistant to counter-evidence. The central evaluative concern is therefore not maximal trust. It is the continued capacity of relational structures to respond to evidence, reopen interpretation, revise stabilized judgments, and preserve meaningful possibilities of exit and renewed relation.

A central problem concerns interpretation. Trust and distrust influence how later events are read. A trusted actor may receive charitable interpretation of negative evidence, while a distrusted actor may find positive evidence recoded as strategic deception. Such dynamics can consolidate interpretive authority and create asymmetries in whose descriptions of public reality acquire weight. Publicity alone does not remove these asymmetries. Status, institutional position, social connection, attention, expertise, and prior reputation can shape the circulation and credibility of competing interpretations.

The field-governance perspective is introduced to examine these distributed and coupled conditions without assuming that every governance problem can be solved through a central intervener or through individually exceptional judgment. A designed governance field, provisionally denoted by ϕ_e , is treated as an intervention within an already evolving relational configuration rather than as an external and omniscient controller. Its effects may couple with trust, distrust, evidence circulation, interpretive authority, status, attention, and other relational conditions. Higher-order effects are therefore relevant: an intervention designed to improve one dimension of public trust can alter other dimensions and can itself become incorporated into existing power structures.

The formal representation remains intentionally incomplete. Expressions based on an action, coupled fields, perturbative expansion, scale, and observation are used to organize questions rather than to claim a physical reduction of social life. The paper leaves open the empirical identification of fields, coupling coefficients, observables, appropriate scales, and conditions under which a field representation offers explanatory advantages over network, game-theoretic, institutional, or interpretive approaches.

The provisional governance concept developed through the paper is public re-cognizability. A public judgment should not be considered adequately open merely because its conclusion is visible. A recognizable public structure preserves the possibility of reconstructing how a judgment emerged, what evidence and interpretations contributed to it, which alternatives were excluded, how authority was distributed, and under what conditions the judgment can be reopened. Re-cognizability is therefore treated

as one possible condition for keeping trust, distrust, and interpretive power historically revisable.

The paper remains exploratory. It does not offer a settled institutional design, a complete theory of public epistemology, a general field theory of governance, or an empirical model ready for estimation. Its narrower aim is to clarify a family of linked problems and to test whether a preliminary field-governance representation can illuminate the formation, stabilization, ambivalence, and revision of public trust.

Responsible Use and Rights Reservation

This section separates requested scholarly conduct from the legal permissions stated on the following page. It records an ethical request for responsible use and then defines the narrower scope of retained legal rights.

The author encourages good-faith discussion, criticism, independent inquiry, and responsible use of the material in this work. Separately from the licence's terms, the author asks users to consider foreseeable harms when adapting or applying the proposed framework. This ethical request leaves the licence's permissions and legally authorized uses unchanged.

The author retains the rights preserved under CC BY-NC 4.0 and may pursue remedies to which the author is legally entitled for breach of the licence or violation of the author's independently applicable rights. Reuse remains independent from authorial endorsement. Third-party rights require authorization from their respective holders where applicable. Copyright exceptions and limitations, including applicable forms of fair use or fair dealing, remain fully available.

Notices

This page consolidates the manuscript's publication status, licence, development disclosure, research-programme relation, and suggested citation.

Status. This working draft records an exploratory stage of an inquiry into public trust, distrust, interpretive power, relational stabilization, and field-governance representation. The paper does not present a complete theory of field governance and does not claim that public relations are literally physical fields. The field and attractor vocabularies are used as preliminary analytical and formal devices whose explanatory limits, empirical interpretation, and normative implications remain subjects of further research. Definitions, formal constructions, section structure, empirical applications, literature coverage, and originality claims remain subject to revision.

Licence. Except where otherwise indicated, copyright 2026 Wanhong Huang. This work is made available under the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0). Subject to its terms, the licence permits sharing and adaptation for noncommercial purposes with appropriate attribution, a link to the licence, an indication of changes, and attribution that preserves the licensor's independence from the reuse. Reuse is governed solely by that licence; the responsible-use request on the preceding page remains separate from its terms. The licence deed and legal-code link are available at <https://creativecommons.org/licenses/by-nc/4.0/>. The licence governs in case of conflict with this summary. Third-party material remains subject to the rights held by its respective rights holders.

Statement on the use of language models. The exploratory discussions and preparation of this paper involved OpenAI's ChatGPT. ChatGPT supported exploratory dialogue, formal reconstruction, source discovery followed by website verification, exact calculation, argumentative criticism, and drafting in $\text{\LaTeX}$. The author selected the research questions, directed and approved the theoretical commitments and epistemic status of the claims, and bears sole responsibility for the manuscript, including its definitions, formal constructions, taxonomy, calculations, arguments, conclusions, and errors. Authorship credit remains with the human author. Cited sources should be independently verified before publication.

1 Introduction

Public trust is commonly treated as a condition to be restored, strengthened, or maintained. Governments seek public confidence, organizations invest in credibility, institutions design transparency and accountability measures, and cooperation is often described as easier when actors trust one another. These practical concerns rest on a substantial literature in which trust has been related to vulnerability and risk, expectations about another actor's conduct, the reduction of social complexity, and judgments of political trustworthiness [26, 25, 16, 21]. Yet a governance problem appears once trust is considered not only as an attitude toward an already given object, but also as a relational condition that changes the subsequent environment in which judgment and interaction occur. Trust can enable interaction that produces further reasons for trust; distrust can reduce interaction and corrective experience in ways that reproduce distrust. Both can influence how later evidence is interpreted, whose account receives credibility, and which relational possibilities remain available.

This paper develops a preliminary field-governance perspective on that problem. It does not propose a general theory of field governance, nor does it claim that public relations are literally physical fields. The narrower purpose is to examine whether field-like and dynamical representations can clarify several features of public trust that become difficult to describe when trust is treated as an isolated individual attitude or as a scalar policy target. The paper is especially concerned with the importance of initial relational conditions, the historical stabilization of trust and distrust, the coupling of trust with interpretive authority and public power, and the possibility that governance interventions themselves alter the relational field in which they operate. The resulting account treats public trust as normatively ambivalent: trust can support cooperation and continuity, but it can also support dependency, collusion, extraction, or interpretive closure; distrust can support verification and critical distance, but it can also become self-reinforcing and resistant to revision.

The discussion therefore shifts the governance objective away from maximizing trust. Its provisional concern is whether public relations preserve the capacity for later evidence, alternative interpretation, institutional review, and relational exit to make a difference. This capacity is developed through the concept of *public re-cognizability*: the possibility of returning to a stabilized public judgment, reconstructing how it was generated, and reopening it when its evidential, interpretive, or relational conditions change. Public re-cognizability is not offered as a complete criterion of legitimate governance. It is introduced as one way to identify a problem that both high-trust and high-distrust environments can share: the loss of pathways by which established interpretations can be reconsidered.

1.1 Public Trust as a Governance Problem

Trust is difficult to govern in part because the mechanisms that make cooperation possible are heterogeneous. Some cooperative relations depend heavily on confidence in another actor's dispositions or intentions. Others rely on institutional reputation, professional status, law, guarantees, monitoring, collateral, verification, or the possibility of sanction. A security mechanism may permit cooperation precisely because the participants do not need to trust one another very much. Conversely, an institutional or symbolic guarantee may transfer trust from a particular actor to a wider order: a person may be accepted because a university, profession, regulator, public office, or other recognized structure stands behind the interaction. These arrangements should not be collapsed into a single substance called trust. The distinction between trust and trustworthiness is already important in political research, while organizational accounts have emphasized the relation between trust, vulnerability, and perceived characteristics of the trustee [21, 26]. Hardin's account likewise places trust within an ongoing relational structure rather than treating it as an undifferentiated moral good [16].

The governance difficulty considered here begins one step later. Mechanisms that initially substitute for trust can also create conditions under which trust later develops. Verification, limited commitments, institutional safeguards, and low-risk cooperation may permit interaction under severe uncertainty. Successful interaction then adds shared history to the relation. A small initial disposition toward cooperation can therefore alter later evidence, and later evidence can alter the disposition in return. In schematic form, if R_t denotes the relational configuration at time t and T_t a trust-related component of that configuration,

then the relevant structure is not merely $R_t \mapsto T_t$. It also includes a return path,

$$R_t \longrightarrow T_t \longrightarrow R_{t+1},$$

where T_t changes the interactions, exposures, interpretations, and opportunities from which R_{t+1} is generated. The same form can arise for distrust. A weak initial distrust condition can produce social distance, reduced direct interaction, defensive conduct, and greater reliance on indirect interpretation; those consequences can then supply further material for distrust. The paper later describes these processes as *attractor-like* while retaining important qualifications about the limits of that language.

The significance of this feedback is that an initially constructed condition need not remain merely superficial. A ceremonial environment, a professional role, a reputation signal, a shared institution, or a deliberately protected first interaction may alter the probability that a relation begins in one direction rather than another. Subsequent history can then stabilize what began as a weak or externally supported condition. The converse also holds. An early cue of suspicion, especially when introduced by an actor with high status or dense social connections, can change how others approach a person or institution. Changed treatment can induce changed responses, and those responses can later be interpreted as evidence supporting the original suspicion. In such cases, an interpretation does not simply describe an independent social reality; through its relational consequences, it participates in producing later evidence.

This makes public trust a governance problem rather than only a measurement problem. A policy that increases reported trust may still generate undesirable dependence or weaken mechanisms of correction. A policy that institutionalizes skepticism may protect accountability while also making cooperation progressively more difficult. The normative valence of trust therefore cannot be inferred from its magnitude alone. Two states may exhibit high trust while jointly producing severe harms toward outsiders; a firm may cultivate trust and subsequently monetize the resulting dependency; a public authority may receive such strong presumptive credibility that counter-evidence becomes increasingly difficult to hear. Conversely, distrust can motivate auditing, independent verification, and resistance to concentrated authority, yet it can also produce a structure in which every favorable event is interpreted as deception and every unfavorable event as confirmation.

The central contrast is consequently not between trust as good and distrust as bad. It is between relational configurations that remain responsive to historical revision and configurations that become self-sealing. Trust can reduce unnecessary suspicion while retaining sensitivity to evidence. Distrust can preserve critical distance while remaining capable of being changed by successful cooperation. Either can also harden into an interpretive structure that increasingly determines in advance what later events are allowed to mean. Public governance must therefore attend not only to the distribution of trust, but also to the dynamics through which trust and distrust alter the future conditions of their own evaluation.

The adjective *public* is used in this paper in a deliberately broader sense than a technical synonym for citizen confidence in government. Political trust remains an important reference point [21], but the present inquiry also concerns trust and distrust among actors, institutions, claims, and interpretive communities when their effects extend into a shared public environment. A relation becomes relevant to the present analysis when its trust dynamics affect public coordination, evidence circulation, participation, reputation, interpretation, or access to common institutions. This usage allows the paper to examine a common structure across political institutions, organizations, public controversies, and other settings without assuming that all such domains are substantively identical.

1.2 Scope and Analytical Orientation

The paper adopts a generative relational orientation. The term *relational* indicates that trust is not treated solely as an intrinsic property of a trustor or trustee. It depends on histories of interaction, institutional mediation, vulnerability, symbolic and affective conditions, expectations, and the positions from which actors encounter one another. The term *generative* adds a temporal claim: once a trust-related configuration emerges, it can contribute causally to the production of later relational conditions. A trusted actor may receive wider discretion, more information, more charitable interpretation, and more opportunities for

cooperation. A distrusted actor may receive less contact, fewer opportunities to produce counter-evidence, more intensive monitoring, and less interpretive charity. These consequences can in turn modify the relation that generated them.

This orientation does not require an ontological claim that trust exists as an independent entity. It is sufficient for the present argument that a configuration interpreted and enacted as trust can acquire stable causal significance within a relational history. The paper therefore distinguishes, at least provisionally, between the truth of a belief about another actor's trustworthiness and the relational reality of trust itself. A belief may be poorly grounded while still reorganizing disclosure, cooperation, resource flows, dependence, or public interpretation. An initially mistaken trust can become historically consequential; an initially weak suspicion can become socially productive. Relational reality in this sense does not confer epistemic correctness or normative legitimacy. It identifies causal participation in the evolving relational system.

The field-governance perspective is introduced at this point as a limited representational strategy. Public trust is rarely shaped by a single dyadic relation. Status, attention, institutional position, evidence access, reputation, network connection, security arrangements, and interpretive authority can be distributed across many actors and locations. A field representation allows these conditions to be considered as jointly shaping the local possibilities of action and interpretation without presuming a single central controller. The paper later introduces a collection of relational fields, provisionally denoted by Φ , and a designed intervention field ϕ_e . The purpose is not to reproduce physical field theory in a social setting, but to give a compact language for distributed influence, coupling, feedback, spatial or network heterogeneity, and the possibility that interventions have effects beyond their nominal target.

The proposed intervention field is therefore internal to the problem it seeks to govern. Once introduced, a transparency rule, review process, evidentiary standard, public forum, verification mechanism, or other intervention can be interpreted, strategically used, resisted, institutionalized, or captured. Its consequences need not remain confined to trust. It may alter attention, status, participation, reputation, evidence circulation, and the relative authority of competing interpretations. This motivates the paper's later use of higher-order coupling and perturbative language. The formalism is intentionally incomplete. Coupling coefficients are not estimated, a universal field content is not specified, and no claim is made that every relevant social mechanism admits a useful field representation. The formal vocabulary is used to expose dependencies that a one-variable account of public trust can obscure.

The same restraint applies to the paper's use of balance. Public balance is not defined as equality of influence among all actors. Expertise, direct experience, institutional responsibility, and evidential access can provide legitimate reasons for unequal weight. Nor is balance treated as a fixed social state. Public influence is multidimensional, and the appearance of balance can change with the observational scale and the variables included. A forum can allocate equal speaking time while leaving agenda-setting, evidentiary access, reputation, and subsequent circulation highly unequal. The more relevant concern is whether asymmetries become capable of closing the processes through which their own authority could later be contested.

This concern leads to public re-cognizability. Transparency usually emphasizes visibility: the public can see a decision, a record, or the evidence that an institution chooses to disclose. Re-cognizability adds a historical and procedural dimension. It concerns whether later participants can reconstruct the formation of a public judgment, distinguish evidence from interpretation, identify the distribution of interpretive authority, recover excluded alternatives where appropriate, and determine the conditions under which the judgment may be reopened. A re-cognizable structure does not require permanent indecision. Stable judgments, trusted institutions, professional authority, and durable cooperation are compatible with it. The relevant condition is that stabilization does not erase the pathways by which revision could become justified.

The paper accordingly treats governance as the design and maintenance of conditions for revisable stabilization rather than as the production of an ideal trust level. This formulation also limits the ambition of the argument. The paper does not derive a complete normative theory of public legitimacy, justice, or democratic participation. It does not determine the correct distribution of expertise or authority in every public domain. It does not supply a universal metric for re-cognizability. Its aim is narrower: to show why the dynamics of trust and distrust can make such questions unavoidable, and to develop a preliminary

field-governance vocabulary with which they can be stated more precisely.

1.3 Structure of the Discussion

Section 2 establishes the conceptual background by distinguishing trust, distrust, and reliance; locating them within institutional and relational conditions; and examining the relation between public trust and interpretive authority. Section 3 turns to the relational formation of trust and distrust. It considers initial conditions, shared experience, symbolic and institutional mediation, and the sense in which trust-related configurations can become relationally real through their continuing causal effects.

Section 4 develops the paper's attractor-like language. It examines positive feedback, path dependence, basin transitions, hysteresis, and relational lock-in while emphasizing that the underlying dynamics can themselves change over time. Section 5 then analyzes the interpretive consequences of stabilization. Trust and distrust are treated as conditions that can change the reading of evidence, privilege particular interpreters, and in limiting cases produce self-sealing interpretive structures. Re-cognizability is introduced there as a counter-concern centered on the continued possibility of revision.

Section 6 extends the analysis from dyadic relations to the public distribution of interpretive power. It considers status, connection, publicity, reputation, exclusion, and institutional mediation, with particular attention to the possibility that unequal interpretive weight becomes self-reinforcing. Section 7 introduces the preliminary field-governance representation. The action, relational fields, intervention field, higher-order coupling, perturbative effects, scale, and observation are presented as analytical devices rather than as a completed formal theory.

Section 8 considers governance under both trust and distrust, including protected cooperation under low-trust conditions, infrastructures through which relational history can accumulate, interventions into distrust attractors, and institutionalized possibilities of rebalancing. Section 9 develops the normative ambivalence of trust and distrust by examining cooperation, dependency, extraction, collusion, externalized harm, critical distance, and interpretive closure. Section 10 draws provisional implications for public governance, focusing on initial relational conditions, interpretive concentration, relational lock-in, revision, exit, and field-level conditions for public trust.

Section 11 states the main limitations of the inquiry, including the limits of field and attractor language, the problem of empirical identification, and the restricted scope of the normative claims. Section 12 concludes by returning to the paper's central proposal: public trust should be governed neither as an isolated psychological quantity nor as an unconditional social good, but as part of a historically evolving relational configuration whose stability must remain compatible with evidence-sensitive revision and public re-cognizability.

2 Conceptual Background

This section establishes the conceptual distinctions required for the later dynamical account. Its purpose is not to settle the philosophy of trust, a literature containing several competing accounts, but to identify a minimal structure adequate to the governance problem developed in this paper. Four points are especially important. First, trust, distrust, and reliance should not be treated as positions on a single undifferentiated scale. Second, the conditions that support cooperation can be institutional and relational even when interpersonal trust is weak. Third, public trust affects the distribution of interpretive authority by changing whose testimony, explanations, and assurances are granted weight. Fourth, an exclusively individualistic representation loses the feedback through which trust-related attitudes alter the relations from which later attitudes are formed. These distinctions prepare the transition from conceptual analysis to the relational and attractor-like dynamics developed in Sections 3 and 4.

2.1 Trust, Distrust, and Reliance

Trust is often associated with favorable expectations concerning another actor, but favorable expectation alone is too broad for the present inquiry. A person may expect a train to arrive, a locked door to remain closed, or an opportunistic contractor to perform because a penalty makes defection costly. Such expectations can support action while involving little of the vulnerability ordinarily associated with trusting another actor. Organizational research has therefore connected trust with a willingness to accept vulnerability under conditions in which another party's conduct matters to the trustor [26]. Philosophical accounts have emphasized different elements of this structure. Baier draws attention to the distinctive vulnerability of trusting relations and to the difference between betrayal and ordinary disappointment [2]; Jones develops an affective account in which trust includes an attitude of optimism about another's goodwill and competence within a domain of interaction [19]; Hardin locates trust within relationally grounded expectations concerning another actor's interests [16]. These accounts are not interchangeable, but each resists reducing trust to a detached probability estimate.

The present paper therefore uses *trust* as a family-resemblance term for relational conditions in which an actor accepts some exposure to another actor, institution, or organized practice while treating favorable conduct as sufficiently expectable to proceed. The formulation is intentionally permissive. It does not assume that every case of trust depends on benevolence, calculative interest, affect, moral commitment, or a single psychological mechanism. The paper instead focuses on a structural consequence shared by many such cases: trust can change what a subject is willing to disclose, delegate, risk, interpret charitably, or leave unverified. In this sense trust matters because it reorganizes the practical and interpretive conditions of subsequent interaction.

Reliance is used more broadly. An actor relies on another person or mechanism when a plan depends on its expected performance, even when that expectation is secured through monitoring, sanctions, redundancy, technical constraint, or other safeguards. Cooperation can consequently persist under conditions of weak trust. Cook, Hardin, and Levi develop this point explicitly by showing how credible commitments, enforcement, reputation, monitoring, and institutional arrangements can support cooperation without requiring generalized trust [8]. This distinction is central to the paper's later governance argument. A mechanism that reduces the amount of trust required for cooperation is not thereby hostile to trust. By lowering the potential cost of early interaction, it can create the conditions under which a shared relational history later becomes possible.

Trust and reliance can therefore overlap while remaining conceptually distinct. If actor *A* cooperates with actor *B* only because a perfectly enforceable contract makes opportunism unprofitable, the immediate cooperation may rest primarily on reliance upon the enforcement structure. If repeated performance under that structure changes *A*'s expectation of *B* across contexts where enforcement is incomplete, the history produced by reliance may contribute to later trust. The direction can also reverse. Established trust may lead actors to remove safeguards, increase discretion, or expand exposure, thereby replacing some institutionally secured reliance with relation-dependent vulnerability. The distinction is thus not a taxonomy of mutually exclusive social states; it identifies different mechanisms through which coordination becomes possible.

Distrust requires equal care. Treating distrust as merely the numerical opposite of trust assumes a single continuum on which more of one necessarily means less of the other. Several accounts resist that simplification. Luhmann treats trust and distrust as distinct ways of managing social complexity rather than as a simple presence and absence of the same state [25]. Lewicki, McAllister, and Bies argue that trust and distrust can coexist within a relationship when actors have differentiated expectations across dimensions and domains [22]. An institution can, for example, be trusted to perform a technical task while distrusted with respect to disclosure, political neutrality, or the use of personal information. A state can be regarded as reliable in maintaining a treaty procedure while remaining distrusted in broader strategic intentions. Public relations are frequently composed of such mixtures.

For the purposes of this paper, distrust is therefore treated as a condition that increases the salience of adverse possibilities and makes protective, verificatory, distancing, or interpretively cautious responses more likely. This working description does not classify all distrust as irrational or socially harmful. Distrust can motivate auditing, checks and balances, evidentiary demands, and other mechanisms that preserve

accountability [8]. The later normative argument depends on this ambivalence. Trust can become dogmatic, and distrust can remain critically productive. Conversely, distrust can become self-sealing, and trust can remain responsive to counter-evidence. The important distinction is accordingly not a moral opposition between trust and distrust, but the different ways in which each organizes exposure, interpretation, and future interaction.

A final distinction concerns *trustworthiness*. Trust is a state or relation involving a trustor; trustworthiness concerns features of the object of trust and the grounds on which favorable expectations are warranted. Political research has repeatedly distinguished levels of trust from judgments about whether institutions or actors merit it [21]. This distinction prevents a recurrent normative error. A society in which citizens trust an institution more is not necessarily a society in which the institution has become more trustworthy. Trust can be generated by reputation, symbolism, social conformity, selective information, or concentrated interpretive authority as well as by improved performance. A field-governance perspective must therefore remain capable of describing an increase in trust while withholding any inference that the increase is epistemically justified or normatively desirable.

2.2 Institutional and Relational Conditions

The distinction between trust and reliance already implies that trust is embedded within a wider architecture of conditions. Legal enforcement, professional accreditation, organizational reputation, insurance, collateral, audit, verification, and procedural review can reduce uncertainty or redistribute risk. These arrangements sometimes substitute for interpersonal trust; at other times they become objects of trust themselves. A person who accepts a professional's judgment may know very little about that individual while relying on licensing, institutional affiliation, malpractice rules, peer review, or the expected consequences of misconduct. What appears phenomenologically as trust in a person can therefore be partly mediated by confidence in a surrounding institutional order.

This mediation matters because institutions alter the exposure structure of interaction. Suppose a relation would be unacceptable when the expected loss from defection is L , but an institutional safeguard reduces the residual loss to L' , with $L' \ll L$. The safeguard does not need to change the actor's estimate of the other's character in order to make cooperation feasible. It changes the consequences of being wrong. In that sense institutions can create a *pre-trust cooperation condition*: a region in which interaction can begin before strong trust is available. The repeated conduct made possible there can subsequently become part of the evidence from which trust or distrust develops. Institutional support and relational trust should therefore not be modeled only as substitutes. They can also be temporally ordered and mutually generative.

The relational dimension extends beyond formal institutions. Trust is produced within histories of exposure, response, interpretation, and memory. The same action can carry different significance depending on what has preceded it. A delayed reply from a long-standing collaborator and the same delay from a previously unreliable actor do not enter identical interpretive contexts. Prior relations supply expectations, default explanations, and thresholds at which protective action becomes reasonable. Hardin's relational account is especially useful here because it resists treating trust as a free-floating attitude detached from the interests and histories of particular actors [16]. The present paper generalizes this relational emphasis beyond dyadic interest to include institutional position, symbolic mediation, public reputation, and the interpretive environment in which conduct becomes meaningful.

Relational conditions can also be materially and affectively mediated. Architecture, ritual, dress, interface design, sensory cues, professional language, spatial distance, and procedural form can affect whether an interaction is experienced as safe, authoritative, familiar, or threatening. The paper does not claim that such cues constitute trust or provide reliable evidence of trustworthiness. Their significance is more limited and more consequential: they can alter the initial conditions under which exposure becomes acceptable. A designed environment can make a subject more willing to enter an interaction, disclose information, accept delay, or interpret ambiguity charitably. Once interaction begins, the resulting history can either correct the initial impression or stabilize it.

This distinction between *conditions of entry* and *grounds of justification* is necessary for avoiding two opposite mistakes. The first treats every designed trust cue as manipulation. Public administration, health

care, conflict mediation, and other practices may legitimately reduce unnecessary fear or ambiguity so that interaction can occur. The second mistake treats a subject's felt safety as evidence that the trusted actor or institution is worthy of the resulting confidence. A calm environment, prestigious title, familiar symbol, or impressive building can affect readiness to engage while saying little about how the institution will behave when interests conflict. Field governance is concerned with the first-order effects of such conditions and with the higher-order histories they permit, while normative evaluation must remain sensitive to the second distinction.

Institutions also distribute the capacity to exit and revise relations. A contract, appeals process, ombuds function, second-opinion mechanism, independent review body, or reversible pilot program can reduce the cost of discovering that initial trust was misplaced. These arrangements matter not because they eliminate vulnerability, but because they prevent vulnerability from becoming irrevocable too early. They can therefore support a form of trust that remains compatible with later correction. Conversely, institutional arrangements can deepen dependence by making exit costly, concentrating information, or tying multiple resources to the same trusted intermediary. The institutional environment is thus part of the dynamics of trust itself rather than a neutral background against which an otherwise psychological attitude operates.

The conceptual consequence is that a public trust relation cannot be adequately characterized by a pair (A, B) alone. At minimum, the relation depends on a surrounding configuration C_t containing relevant institutions, safeguards, reputational signals, audiences, histories, and available alternatives. A schematic representation can therefore write trust-related conduct as

$$a_t = F(A, B, H_t, C_t),$$

where H_t denotes the relational history available at time t . The function F is not proposed as an estimable universal law. It marks a dependency that later sections will elaborate: changes in C_t can change interaction even when neither actor's intrinsic properties have changed, and the interaction generated at t becomes part of H_{t+1} and potentially of C_{t+1} itself.

2.3 Public Trust and Interpretive Authority

Trust affects public relations not only by changing willingness to cooperate, but also by changing the weight assigned to interpretation. When an institution is trusted, its explanations can receive presumptive credibility; when an actor is distrusted, even accurate statements can face a higher evidentiary threshold. This does not make trust irrational. Public life depends extensively on epistemic dependence. No participant can personally reproduce every scientific analysis, audit every administrative record, observe every diplomatic negotiation, or independently verify every technical judgment. Some distribution of epistemic authority is therefore unavoidable. The governance problem arises when the distribution of trust and distrust begins to determine which interpretations remain contestable.

A useful starting point is to distinguish *interpretive authority* from truth. Interpretive authority is the socially effective capacity to make an account of events count: to frame relevant categories, identify salient evidence, assign causal significance, or define which uncertainties deserve attention. Such authority can rest on expertise, direct experience, institutional role, reputation, social status, network position, or public visibility. These bases of authority are heterogeneous. A specialist may possess strong domain knowledge but limited public reach; a highly connected actor may possess little direct evidence but considerable agenda-setting capacity. Treating public judgment as though interpretations entered a neutral arena with equal effective weight obscures these differences.

The literature on epistemic injustice supplies an adjacent vocabulary for this concern. Fricker analyzes how social power and prejudice can affect the credibility granted to speakers and can impair participation in shared interpretive resources [10]. The present paper does not equate every trust asymmetry with epistemic injustice, nor does it adopt testimonial or hermeneutical injustice as a complete model of public trust. The relevance is structural: credibility judgments are socially situated, and unequal credibility can become part of the conditions under which subsequent knowledge claims are heard. A field-governance account therefore asks how trust, status, institutional position, and interpretive access become coupled over time.

The coupling can operate in several directions. Prior trust can increase the default weight of an actor's account. Increased interpretive weight can shape the behavior of others, for example by changing whom they approach, exclude, monitor, or consult. Those altered relations can generate new observations, which are then interpreted within the original trust configuration. Conversely, distrust can reduce direct contact and increase dependence on indirect reports, leaving fewer opportunities for corrective experience. In both cases, trust or distrust participates in producing the evidentiary environment within which it will later be assessed.

This structure is especially important in public settings because visibility and connectivity are uneven. If actor i offers interpretation I_i of an event, the public effect of that interpretation depends on more than its propositional content. It can be represented schematically as

$$E(I_i) = W_i I_i,$$

where W_i stands for a multidimensional and context-dependent effective weight incorporating such factors as status, connection, institutional position, audience access, and prior credibility. The expression is not a numerical model of persuasion. It serves to distinguish formal participation from effective interpretive influence. A public forum can permit every participant to speak while leaving W_i highly unequal.

Unequal interpretive weight is not by itself a defect. Expertise, evidential access, and responsibility can justify substantial asymmetry. A public health specialist and an uninformed observer need not receive identical weight on a technical question. The governance concern is instead the possibility of *self-amplifying interpretive privilege*. If prior status increases credibility, credibility increases circulation, circulation increases institutional access, and institutional access produces further opportunities to define the evidence, then an initially justified asymmetry can become progressively insulated from the conditions that justified it. The same mechanism can attach to distrust: an actor initially regarded as unreliable can receive fewer opportunities to produce public evidence, making the original judgment increasingly difficult to revise.

Publicity alone does not solve this problem. Making a controversy visible can expose claims to scrutiny, but it can also magnify differences in attention, reputation, and network position. Nor does majority acceptance establish truth. The more defensible function of public governance is procedural and relational: to preserve routes through which evidence can enter, affected subjects can respond, competing interpretations can be articulated, conflicts of interest can be identified, and established judgments can later be reopened. This is the context in which the paper's concept of public re-cognizability becomes relevant. Re-cognizability concerns the recoverability of a judgment's generative history and the continued availability of conditions under which its interpretation can change.

The distinction also clarifies why trust can become normatively ambivalent even when it begins as a reasonable response to expertise or past performance. Trust economizes attention and reduces the need for constant verification. Precisely for that reason, it can concentrate interpretive discretion. A trusted institution may be granted greater latitude to select which evidence to disclose, which categories to use, and which anomalies require explanation. Such discretion can be efficient and legitimate, yet it also creates a pathway by which trust can weaken the mechanisms that would reveal declining trustworthiness. The later sections therefore treat interpretive closure not as an accidental corruption of trust from outside, but as one possible trajectory produced by the same delegation and reduced monitoring that make trust practically valuable.

2.4 Limits of Individualistic Accounts

Individual-level accounts of trust are indispensable for many purposes. Surveys of confidence, experiments on cooperative behavior, psychological models of risk, and formal representations of belief can identify important differences among persons and contexts. The limitation arises when public trust is modeled exclusively as an attribute stored within individuals and later aggregated. Such a representation can describe a distribution of attitudes at time t , but it can miss how the distribution changes the relational environment from which attitudes at time $t + 1$ emerge.

Consider a population in which each actor has a trust value $T_i(t)$. A conventional aggregate might be written as

$$\bar{T}(t) = \frac{1}{N} \sum_{i=1}^N T_i(t).$$

The quantity can be empirically useful, but two systems with the same $\bar{T}(t)$ can have very different relational structures. In one, trust may be broadly distributed across weak ties and supported by robust verification. In another, the same average may result from intense trust concentrated around a small number of central actors while large portions of the population remain excluded or distrustful. The first arrangement may preserve multiple pathways of correction; the second may be vulnerable to interpretive monopoly or cascading failure. The mean does not identify these differences.

The same difficulty applies to temporal change. If an actor becomes more trusted, others may disclose more information, grant more discretion, reduce monitoring, and offer more favorable interpretations. These consequences change the actor's opportunities and the evidence available to observers. Trust is therefore not merely an output of prior evidence; it can become an input into the process that generates future evidence. A minimal representation requires an endogenous loop,

$$T_t \longrightarrow R_{t+1} \longrightarrow E_{t+1} \longrightarrow T_{t+1},$$

where R denotes relational configuration and E the evidence or experience produced within it. Distrust can enter an analogous loop. The later attractor-like analysis begins from this endogeneity.

An individualistic model also tends to obscure the difference between changing an attitude and changing the conditions under which the attitude is formed. A public campaign can attempt to persuade citizens directly, but an institutional reform can instead alter auditability, participation, visibility, exit costs, or the distribution of interpretive authority. Both may change reported trust, yet they act at different levels. The field-governance perspective is motivated by the second class of intervention: governance can modify the environment in which trust and distrust are generated rather than simply targeting the subjects who hold them.

A further limitation concerns the unit of trust. Statements such as “the public trusts institution X ” compress several distinct relations. Citizens may trust an institution's technical competence while distrusting its political neutrality; trust frontline personnel while distrusting senior leadership; rely on its procedures while rejecting its public explanations; or trust it under ordinary conditions while expecting opportunism during crisis. Lewicki, McAllister, and Bies's multidimensional treatment of trust and distrust cautions against forcing these configurations onto a single line [22]. For public governance, domain specificity is not a secondary detail. Different components of trust can generate different kinds of exposure and therefore different kinds of systemic risk.

The alternative proposed here is not to discard individual attitudes but to embed them. Trust-related states are treated as components of a larger evolving configuration that includes relational history, institutional scaffolding, public interpretation, and distributions of power. Symbolically, the object of analysis becomes

$$\mathcal{R}_t = (X_t, H_t, C_t, I_t, P_t),$$

where X_t contains actor-level states, H_t relational histories, C_t institutional and material conditions, I_t interpretive structures, and P_t relevant power asymmetries. The components are provisional rather than exhaustive. Their function is to make one methodological commitment explicit: public trust cannot be understood by holding the environment fixed when trust itself helps to change that environment.

This commitment defines the boundary of the conceptual background developed in this section. Trust is distinguished from reliance and trustworthiness; distrust is treated as potentially coexisting with trust and as capable of both critical and self-sealing forms; institutions are understood as both substitutes for and generators of relational history; and public trust is linked to the distribution of interpretive authority. Section 3 now turns from these distinctions to formation. It examines how initial conditions, shared experience, symbolic and institutional mediation, and repeated interaction can produce trust and distrust as historically consequential relational structures.

3 Relational Formation of Trust and Distrust

This section develops the transition from conceptual distinctions to formation processes. The central claim is modest: trust and distrust are not adequately represented as attitudes that appear only after an actor has completed an isolated assessment of another actor. They are also shaped by the conditions under which a relation becomes available, by histories produced through repeated interaction, and by symbolic, sensory, and institutional environments that alter what forms of exposure, interpretation, and cooperation are initially possible. The discussion therefore proceeds from relational entry to historical accumulation, then to mediation by public and embodied conditions, and finally to the sense in which trust and distrust can become relationally real even when their underlying judgments remain uncertain or mistaken. The stronger dynamical claim that these structures may become attractor-like is reserved for Section 4.

3.1 Initial Conditions and Relational Entry

A relation rarely begins from a neutral state. Before two actors have accumulated substantial direct experience, their first interaction is already conditioned by prior reputation, institutional affiliation, social status, third-party testimony, spatial setting, perceived risk, available safeguards, and the expectations associated with the roles they occupy. These factors need not amount to trust. Their immediate significance is that they alter the threshold at which an actor is willing to enter a relation at all.

Let $\mathcal{R}_0$ denote the relational configuration at the beginning of an interaction. A schematic decomposition is

$$\mathcal{R}_0 = (X_0, C_0, S_0, G_0, V_0),$$

where X_0 collects relevant actor-level states, C_0 material and institutional conditions, S_0 symbolic and reputational cues, G_0 the initial network or role structure, and V_0 the distribution of vulnerability. The notation is not intended as a complete ontology. Its purpose is to make explicit that an encounter begins within a structured environment rather than from an informational vacuum.

This distinction is important because the conditions that permit relational entry can be mistaken for trust itself. A prestigious title, institutional affiliation, contractual safeguard, visible security mechanism, or recommendation from a respected intermediary may make interaction possible even when the actor has little direct confidence in the counterparty. Conversely, the absence of such supports can prevent interaction between actors who might, after sufficient contact, have developed a stable cooperative relation. Initial conditions therefore shape the set of trajectories that can begin without determining the eventual quality of the relation.

The distinction can be expressed by separating an initial willingness to enter a relation, E_0 , from a later trust configuration, T_t . In the simplest schematic form,

$$E_0 = F(\mathcal{R}_0),$$

while

$$T_t = G(\mathcal{R}_0, H_t, I_t),$$

where H_t denotes the history accumulated after entry and I_t the interpretive structures through which that history is understood. The separation prevents a strong initial cue from being read as proof of trustworthiness. A favorable initial environment can lower the cost of first contact; it does not guarantee how the relationship will evolve.

This framing also gives a limited role to safeguards. Monitoring, deposits, guarantees, staged commitments, and reversible forms of cooperation can reduce the potential loss associated with early interaction. As argued in Section 2, these mechanisms can support cooperation without presupposing trust [8]. Their relevance to trust formation lies in the history they make possible. If actors can enter low-risk cooperation under conditions of weak trust, successful interaction may generate evidence and experience that were unavailable before the relationship began. Institutional security can therefore operate as a precondition for relational learning even when it is not itself a form of interpersonal trust.

The same logic applies to distrust. A relation can begin under a weak adverse prior created by warning, reputation, role expectation, or inherited intergroup narrative. Such an initial condition need not contain enough evidence to justify a settled judgment. It may nevertheless alter subsequent behavior by increasing monitoring, reducing disclosure, narrowing contact, or encouraging more defensive interpretations. This makes initial distrust consequential before it becomes deeply established. A small asymmetry at relational entry can change which experiences are subsequently produced and which corrective experiences never occur.

For public governance, the implication is that initial conditions deserve attention independently of final judgments. A governance arrangement that focuses only on whether trust is eventually high or low can miss the earlier mechanisms that determine who receives opportunities for direct interaction, whose assurances are treated as sufficient to begin cooperation, who bears the cost of first vulnerability, and which actors are excluded before a shared history can form. These conditions are especially important in public settings because reputational and institutional cues are unevenly distributed. Actors with greater status or connectivity can enter interactions with inherited credibility, while others must generate credibility under more restrictive conditions.

3.2 Shared Experience and Historical Accumulation

Once interaction begins, trust and distrust can become increasingly historical. A relation that has persisted through repeated encounters contains information that cannot be reduced to a single present judgment. Actors remember whether commitments were kept, whether vulnerabilities were exploited, whether errors were acknowledged, whether uncertainty was disclosed, and how the other party behaved when incentives changed. The resulting history provides material from which later expectations are formed.

Let

$$H_t = \{e_1, e_2, \dots, e_t\}$$

denote the sequence of interactionally relevant events available by time t . A simple learning model might treat trust as an update from H_t to a posterior expectation. Such a representation captures part of the process, but relational history is richer than a list of independent observations. The meaning of e_k can depend on earlier events, role expectations, power asymmetries, and the interpretive stance already present when the event occurs. The same delayed response can be read as negligence, caution, overload, or respect for procedure depending on the relation in which it appears.

Repeated interaction therefore produces two forms of accumulation at once. It produces additional events, and it modifies the interpretive background through which later events become evidence. Hardin's relational account emphasizes that trust concerns a situated relation between a trustor and trustee rather than a free-standing property of the trustee [16]. Repeated-interaction research similarly shows that reputation and observed prior conduct can alter cooperation in trust settings, although experimental trust games capture only a restricted part of the social process. The point needed here is narrower: later interaction takes place in a history that earlier interaction helped create.

This makes successful cooperation potentially generative. Suppose actor A accepts a limited vulnerability to B , B responds in a manner compatible with A 's expectation, and the result becomes part of the relation's shared history. At the next interaction, A does not face exactly the original decision. The relational state has changed. The history can support greater openness, a wider scope of delegation, or a more tolerant interpretation of ambiguity. In schematic form,

$$T_t \longrightarrow A_t \longrightarrow H_{t+1} \longrightarrow T_{t+1},$$

where A_t represents trust-conditioned action. The expression does not imply monotonic increase. A later breach, contextual change, or reinterpretation of earlier events can reverse the direction. Its purpose is to show that trust-related action participates in producing the history from which later trust is evaluated.

Distrust can accumulate through an analogous process. Increased caution can reduce direct contact, limit disclosure, and shift information acquisition toward indirect sources. Those changes can leave fewer opportunities for favorable experience and can increase the salience of ambiguous or adverse evidence.

The resulting history is partly endogenous to the distrust that preceded it. For this reason, a later claim that distrust is “supported by experience” may be correct at one level while still requiring investigation into how that experience was generated. The history of a relation is evidence, but it is also an outcome of earlier relational conditions.

Historical accumulation does not imply that actors must remain bound to the past. New evidence can revise a relation, and institutional intervention can create encounters that would not otherwise occur. The important point is that revision begins from an already structured history. This creates asymmetry between formation and repair. A single favorable interaction may be insufficient to undo a long distrust history, just as one ambiguous failure may be absorbed within a deeply established trust relation. Section 4 develops this asymmetry through the language of path dependence, hysteresis, and attractor-like stabilization.

Shared history also complicates public judgment. Outsiders may observe only a small portion of a relation, while participants possess accumulated experience unavailable to the public. This can justify some difference in interpretive weight. At the same time, appeals to private history can become difficult to scrutinize and can shield established relations from external evidence. The governance problem is therefore not to replace situated knowledge with public abstraction, but to preserve channels through which relational history can be articulated, compared with other evidence, and revised when necessary.

3.3 Symbolic, Sensory, and Institutional Mediation

Trust formation is mediated by more than explicit propositions. People encounter institutions and other actors through architecture, clothing, titles, procedural rituals, language, visual design, sound, smell, bodily proximity, and other features of the environment. These features can affect affective orientation, perceived familiarity, perceived professionalism, and willingness to approach. They should not be treated as direct measures of trustworthiness, but neither are they irrelevant simply because they are non-propositional.

The philosophical literature already provides reasons to resist an exclusively calculative picture of trust. Jones’s account emphasizes trust as an affective attitude rather than a detached estimate alone [19]. At an empirical level, experimental work has reported that environmental olfactory cues can alter behavior in trust-game settings. Sellaro and colleagues, for example, found higher transfers in a trust game under lavender aroma than under peppermint or a no-scent control condition [31]. Such findings should be interpreted cautiously: laboratory measures of transfer do not establish a general law of social trust, and an induced change in trusting behavior says nothing by itself about the actual trustworthiness of the target. Their relevance here is more limited. They demonstrate that conditions conventionally treated as peripheral to rational assessment can enter the formation of trust-related behavior.

This observation permits a useful distinction between a *trust cue* and a *trust ground*. A cue is a condition that changes an actor’s readiness to approach, expose, delegate, or interpret favorably. A ground is a reason that bears on whether the target merits trust in the relevant domain. The two can overlap, but they need not. Institutional symbols, prestigious surroundings, familiar rituals, or sensory comfort can make a relation easier to enter without providing strong evidence that the counterparty is honest, competent, or benevolent. Confusing cues with grounds creates an obvious pathway for manipulation.

The distinction also applies to institutional mediation. Institutions can supply symbols of legitimacy while simultaneously providing substantive safeguards. A professional title may operate as a symbolic cue, but a licensing regime can also encode training requirements, disciplinary procedures, and traceable responsibility. A government seal may function rhetorically, while the administrative procedure behind it may generate records, review rights, and enforceable duties. The governance problem is therefore not solved by dismissing symbols as superficial. Symbolic and procedural layers can be coupled, and public trust can depend on both.

This coupling helps explain why designed environments can have durable relational consequences. A favorable setting can increase the probability that an initial interaction occurs. That interaction can then generate genuine shared history. Over time, the relation may become sustained by repeated conduct rather than by the cue that helped it begin. Conversely, a hostile or stigmatizing environment can suppress contact, generate defensive behavior, and reduce access to corrective evidence. An initially weak cue can therefore influence later relation formation by changing which trajectories become available.

The same mechanism makes public trust vulnerable to capitalization and strategic design. Commercial actors, political institutions, organizations, and movements can invest in environments that increase familiarity, legitimacy, prestige, or safety. Such investment is not inherently wrongful. Hospitals, courts, schools, and public agencies often need environments that reduce unnecessary fear and make participation possible. The normative issue arises when designed trust cues become systematically detached from trustworthiness, when they obscure conflicts of interest, or when the relation produced by them becomes difficult to reassess. Section 9 returns to this ambivalence in the context of dependency, extraction, collusion, and externalized harm.

A field-governance perspective is useful here because symbolic, sensory, and institutional conditions need not be controlled by a single actor. They can be distributed across architecture, organizational practice, media representation, role expectations, and repeated public performance. Their combined effect can alter the local conditions of interaction without any participant possessing a complete view of the process. Treating these conditions as part of relational formation therefore avoids a false choice between purely individual psychology and centralized manipulation.

3.4 Trust and Distrust as Relationally Real Structures

The preceding analysis raises a difficult ontological and epistemic question. If trust can be initiated by symbolic cues, affective conditions, partial evidence, institutional assurances, and historically contingent interpretation, in what sense can it be said to be real? The present paper does not answer by positing a hidden essence of trust that must be independently located inside the subject. It instead adopts a more limited criterion appropriate to the generative relational framework: a trust or distrust configuration can become relationally real when it produces stable and consequential differences in the evolving relation from which it emerges.

Let the relational configuration evolve according to

$$\mathcal{R}_{t+1} = F(\mathcal{R}_t, T_t, D_t, C_t),$$

where T_t and D_t denote trust- and distrust-related conditions and C_t other relevant contextual variables. Trust is relationally consequential when perturbing T_t while holding the comparison meaningful would alter subsequent relational trajectories. Schematically,

$$F(\mathcal{R}_t, T_t + \delta T, D_t, C_t) \neq F(\mathcal{R}_t, T_t, D_t, C_t)$$

for relevant perturbations δT . The expression is a conceptual criterion rather than an empirical identification strategy. It states that trust matters as part of the relation if its presence changes disclosure, delegation, monitoring, interpretation, cooperation, or other future relational conditions in a persistent way.

This criterion separates relational reality from epistemic correctness. Actor A can trust actor B on inadequate grounds. B can later prove untrustworthy. The original trust may nevertheless have been relationally real if it changed what A disclosed, what authority B acquired, how resources flowed, or which opportunities became available. Likewise, a distrust judgment can be mistaken while still reorganizing contact, reputation, and public interpretation. A relationally real structure can therefore be epistemically defective.

Relational reality is also distinct from normative legitimacy. A stable trust relation can support care, cooperation, and public coordination, but it can also facilitate collusion, dependency, exploitation, or exclusion. A stable distrust relation can protect actors from genuine danger, but it can also sustain stigma, defensive escalation, or interpretive closure. The fact that a structure persists and has causal force does not establish that it ought to persist. This separation is necessary if the framework is to avoid deriving normative approval from social stability.

A third distinction concerns ontological commitment. To say that trust is relationally real does not require claiming that trust exists as an independently bounded substance. The framework suspends that stronger question. What it affirms is that certain configurations named as trust can become sufficiently stable in a generative relational system that removing them from the analysis would obscure real differences in

the system's trajectory. This is analogous to the paper's treatment of public interpretation and institutional authority: their reality is approached through the differences they produce within historically evolving relations rather than through a claim about ultimate metaphysical substance.

This position also explains why beliefs about trust can help produce the conditions that later sustain them. If A treats B as trustworthy, A may disclose more, delegate more, and interpret ambiguity more charitably. These actions can give B opportunities to reciprocate, creating a history that was unavailable before A acted on the belief. The resulting relation may eventually be sustained by accumulated experience even if its initial conditions were thin. The reverse is equally possible for distrust. In both cases, an interpreted condition can enter the causal organization of the relation and thereby contribute to its own historical stabilization.

The concept of relational reality should therefore be read as a claim about generative consequence rather than truth by social agreement. Public acceptance does not make every proposition factually correct, and repeated belief does not transform false evidence into true evidence. What can become real is the relational structure produced through acting on those beliefs. This distinction is especially important for public governance, where mistaken interpretations can generate durable distributions of authority, exclusion, access, and vulnerability. Governance must address those consequences without confusing their social reality with the truth of the interpretations that produced them.

The analysis of formation can now be carried one step further. If trust and distrust alter the relations that generate later evidence and interpretation, then repeated feedback can stabilize particular trajectories. Section 4 examines this possibility through attractor-like language, while retaining the caution that social relations are historically adaptive and that the relevant dynamical landscape can itself change through interaction and governance.

4 Stabilization and Attractor-Like Dynamics

The preceding section treated trust and distrust as relationally consequential structures that can be generated from initial conditions, accumulated histories, and mediated environments. This section examines a further possibility: once such structures begin to influence later interaction, they can participate in feedback processes that make some relational trajectories progressively easier to reproduce. Attractor language is introduced for this limited purpose. It provides a way to describe persistence, path dependence, threshold-like transition, and resistance to small perturbations without presuming that social relations possess a fixed phase space or a time-invariant vector field. The analysis therefore proceeds from positive feedback in trust and distrust formation to historical path dependence, then to transitions between relational regimes, and finally to hysteresis and lock-in. The formal expressions are heuristic representations intended to clarify structural relations rather than to identify a complete dynamical law of public trust.

Let a public relational state at time t be represented schematically by

$$z_t = (T_t, D_t, I_t, A_t, V_t, G_t, C_t),$$

where T_t and D_t denote trust- and distrust-related conditions, I_t interpretive structures, A_t distributions of authority, V_t vulnerability, G_t relevant relational topology, and C_t other contextual conditions. A minimal evolutionary representation is

$$z_{t+1} = F_t(z_t, u_t),$$

where u_t includes events, interventions, and other inputs. The time index on F_t is important. In a generative relational setting, interaction can alter not only the state of the relation but also the mechanisms through which later states are produced. Accordingly, the evolution of the transition structure can be written schematically as

$$F_{t+1} = \Psi(F_t, H_{t+1}),$$

with H_{t+1} denoting the history accumulated through the new interaction. This historical mutability distinguishes the present use of attractor language from a literal importation of fixed dynamical systems into public life.

4.1 Positive Feedback in Trust Formation

Trust can become self-reinforcing when trusting conduct changes the conditions under which later evidence is produced. An actor who regards another as sufficiently trustworthy may disclose more information, delegate greater discretion, reduce defensive monitoring, tolerate limited ambiguity, or accept a degree of vulnerability that would otherwise be avoided. These actions create opportunities for reciprocal conduct. If the counterparty responds in ways interpreted as competent, reliable, or non-exploitative, the resulting experience can strengthen the very trust that made the interaction possible.

A simplified feedback sequence is

$$T_t \longrightarrow O_t \longrightarrow R_t^+ \longrightarrow H_{t+1} \longrightarrow T_{t+1},$$

where O_t denotes trust-enabled openness and R_t^+ a positively interpreted response. The sequence should not be read as deterministic. Openness can be exploited, reciprocity can fail, and favorable conduct can be misinterpreted. Its significance is that trust can alter the probability distribution of later experiences rather than merely summarize experiences already obtained.

This mechanism helps explain why initial trust conditions can matter disproportionately even when they are epistemically thin. A small increase in willingness to enter a relation may permit an interaction that would otherwise never occur. Successful interaction then contributes direct relational history, and direct history can gradually replace weak initial cues as the principal basis of the relation. What began as institutionally supported reliance, reputational confidence, or a favorable sensory and symbolic environment can therefore become historically grounded trust if later conduct repeatedly supports it. The trajectory is generated through the relation rather than fully contained in its starting point.

The relevant feedback is also interpretive. Existing trust can increase the range of benign or non-hostile explanations considered plausible when conduct is ambiguous. This tolerance can protect a relation from being destroyed by every small error. In a stable friendship, organization, or public institution, a single delay, misunderstanding, or procedural failure need not erase an accumulated history. Trust therefore functions partly as a temporal buffer that permits local disturbances to be absorbed without immediate regime change. This buffering capacity is one reason stable trust can be socially productive: continuous cooperation would be difficult if every ambiguous event reset the relationship to an initial condition.

The same mechanism creates a normative risk that later sections examine in detail. A charitable interpretive prior can become so strong that disconfirming evidence no longer produces meaningful revision. The distinction between stabilization and closure is therefore essential. Positive feedback in trust formation is compatible with healthy resilience when the relation remains responsive to sufficiently strong counterevidence. It becomes problematic when the feedback process systematically insulates the trusted actor, institution, or interpretation from correction.

For present purposes, an attractor-like trust regime can be represented as a region $\mathcal{A}_T$ in which a range of nearby relational states tend, under ordinary interaction, to remain within or return toward a trust-supporting configuration. Let $\mathcal{B}_T$ denote the corresponding basin-like region. The claim is not that $\mathcal{A}_T$ is a mathematically identified invariant set. It is that once the relation enters some historically stabilized region, small perturbations need not change its qualitative organization:

$$z_t \in \mathcal{B}_T, \quad \|\delta z_t\| < \varepsilon \quad \Rightarrow \quad z_{t+k} \text{ remains trust-supporting for relevant } k.$$

This formulation shifts attention from the momentary level of trust to the resilience of the relational configuration that sustains it.

4.2 Positive Feedback in Distrust Formation

Distrust can display an analogous feedback structure. A prior of suspicion can increase monitoring, reduce disclosure, narrow contact, encourage defensive preparation, and raise the threshold at which ambiguous conduct is interpreted favorably. Those responses alter the behavior of the other party and the kinds of evidence that become available. Reduced contact can eliminate opportunities for corrective experience;

defensive conduct can elicit counter-defensiveness; and the resulting tension can be read as confirmation of the original suspicion.

A schematic sequence is

$$D_t \longrightarrow W_t \longrightarrow E_t^- \longrightarrow I_{t+1}^D \longrightarrow D_{t+1},$$

where W_t denotes withdrawal, defensive monitoring, or constrained engagement, E_t^- the resulting adverse or ambiguous experience, and I_{t+1}^D a distrust-weighted interpretation. Again, the sequence is probabilistic rather than necessary. Distrust can be corrected by compelling evidence, direct experience, institutional safeguards, or third-party intervention. The important point is that distrust can participate in producing the environment from which its later evidence is drawn.

This mechanism is especially consequential in public and organizational settings. A weak initial warning about a person, group, or institution can reduce direct interaction. Reduced interaction leaves later participants more dependent on reputation, hearsay, inherited categorization, or the judgments of central network actors. The target, encountering distance or suspicion, may become defensive or less communicative. Such behavior can then be treated as further evidence that the initial warning was justified. A small interpretive asymmetry can therefore be amplified through changes in relational behavior.

Threshold models of collective behavior provide a useful adjacent insight. Granovetter shows that aggregate outcomes can depend strongly on the distribution of individual thresholds and that groups with similar average preferences can generate very different collective results [14]. Public distrust need not follow the same binary model, but the threshold intuition is relevant: once enough actors withdraw, repeat a warning, or treat suspicion as the prudent default, the cost of maintaining an alternative interpretation can rise for others. What appears *ex post* as widespread agreement may therefore be partly an emergent consequence of sequential response rather than an independent convergence of judgment.

Distrust can also become interpretively self-protective. Positive conduct may be discounted as strategic performance, while negative conduct is treated as diagnostic. In such a configuration, different observations can be mapped toward the same conclusion. The problem is not simply that the conclusion is negative. A negative judgment can be warranted. The structural problem arises when the mechanism of updating becomes sufficiently asymmetric that the judgment loses a credible pathway of revision. Section 5 develops this problem as interpretive closure; here it is sufficient to note that such closure deepens the basin-like stability of distrust.

Research on public trust in risk management has long emphasized that trust and distrust can exhibit history-sensitive dynamics. Slovic's analysis of risk controversies describes processes through which technological and social conditions can systematically erode trust [33]. Subsequent empirical work has also shown that proposed asymmetries between trust-building and trust-destroying information vary by context rather than constituting a universal law. Evidence from tourism risk, for example, found symmetry in settings where an earlier asymmetry hypothesis predicted stronger effects for negative information [9]. This variation supports a cautious formulation: distrust attractors are plausible relational configurations, but their depth, responsiveness, and transition thresholds depend on the domain, prior history, and informational environment.

4.3 Path Dependence and Relational History

Feedback becomes especially important when it produces path dependence. In its restrained sense, path dependence means that the present relational state cannot be understood adequately from current inputs alone because timing, sequence, and earlier contingencies have altered the options and response structures available later. Pierson's account of increasing returns in political development emphasizes precisely these features: small or contingent early events can have large consequences, timing and sequence matter, and established paths can become costly to reverse [30]. The present argument adapts this logic to trust and distrust while avoiding the stronger claim that all trust relations follow increasing-returns dynamics.

Let two relational histories H_t and H'_t lead to superficially similar present observations e_t . A memoryless representation would imply that the same observation generates approximately the same update. A

path-dependent representation permits

$$U(e_t | H_t) \neq U(e_t | H'_t),$$

where U denotes the updating process. A missed commitment after ten years of reliable cooperation may be interpreted differently from the same missed commitment at the first interaction. Likewise, an ambiguous conciliatory gesture after a history of exploitation may carry less weight than an identical gesture in a relation with little adverse history. Relational history changes the evidential meaning of present events.

Path dependence also operates by changing available action. Trust can produce deeper delegation, shared infrastructure, mutual dependence, common routines, and denser information channels. Distrust can produce separation, duplicated capacities, screening procedures, security measures, and alternative alliances. These structures remain after the original events that generated them have passed. The present state therefore contains sedimented history. Public trust is not only remembered psychologically; it can be embedded materially and institutionally in the architecture through which later interaction occurs.

This point complicates simple causal narratives about public confidence. Suppose an institution introduces a reform and public trust subsequently rises. The observed change may depend on whether the reform enters a relation already characterized by residual confidence, accumulated suspicion, prior failed reforms, or credible intermediaries. The same formal intervention can therefore produce different trajectories in different historical regions of the state space. A field-governance analysis must attend to these inherited conditions rather than assuming that an intervention has a context-independent effect.

Path dependence also clarifies why early governance can matter without implying that early conditions determine everything. The existence of a path does not eliminate agency, intervention, or novelty. It changes their cost and likely effect. A relation with a deep history of distrust may require stronger or differently structured intervention than one in which distrust is recent and weak. Conversely, a highly trusted institution may absorb isolated failures that would severely damage a newly established institution. Historical stabilization therefore generates unequal sensitivities to similar perturbations.

The generative relational perspective adds one further complication. The path can change the landscape through which the path proceeds. Repeated interaction may alter the transition mechanism F_t itself by changing roles, norms, channels, resources, and interpretive habits. Thus the relevant history is not merely a trajectory through a fixed environment. It is partly a history of environment formation:

$$z_t \longrightarrow H_{t+1} \longrightarrow F_{t+1} \longrightarrow z_{t+1}.$$

This recursive structure is central to the present paper's use of attractor language. Public trust and distrust can become stable partly because their histories reshape the conditions of their own reproduction.

4.4 Basin Transitions and External Intervention

If trust and distrust can form basin-like regimes, governance must consider how transitions between them occur. A small perturbation within a deep basin may produce only temporary deviation. A larger perturbation, a sequence of disturbances, or a change in the transition structure can move the relation across a boundary into a qualitatively different regime. The relevant intervention therefore need not aim directly at the final level of trust. It may instead alter the conditions under which a different trajectory becomes reachable.

Consider a distrust-supporting region $\mathcal{B}_D$. An intervention u_t can operate in at least three analytically distinct ways. It can change the current relational state,

$$z_t \mapsto z_t + \delta z,$$

it can change the transition mechanism,

$$F_t \mapsto F_t + \delta F,$$

or it can change the topology of available relations by adding or restoring channels through which evidence, interaction, or review can occur. These mechanisms correspond roughly to event intervention, rule or institutional intervention, and relational restructuring. In practice they can occur together.

The distinction matters because state-level persuasion may be ineffective when the mechanisms that reproduce distrust remain unchanged. Asking actors to “trust more” changes neither the cost of vulnerability nor the evidential environment. By contrast, staged cooperation, independent verification, reversible commitments, third-party observation, or protected channels of contact can reduce the consequences of early exposure and generate new interaction histories. Such mechanisms do not require trust to precede cooperation. Their more modest function is to make trajectories available that a low-trust configuration would otherwise suppress.

The same logic applies to public reputational dynamics. Once an interpretation has become dominant, a target’s own denial may have little effect because the denial is processed through the same distrust-weighted interpretive mechanism. A transition may require evidence or relations that are not fully governed by the local interpretive regime: independent review, a trusted bridge actor, a procedurally protected hearing, or access to records that can be evaluated under different evidential conditions. “External” here is relative rather than absolute. No intervention stands outside the wider social system. The relevant feature is partial independence from the specific feedback loop that is being interrupted.

This point also limits the romantic appeal of charismatic or central correctors. A respected teacher, public figure, mediator, or institution can sometimes supply a perturbation large enough to reopen a closed trajectory. Yet a governance model that depends on the accidental presence of such an actor reproduces unequal access to rescue. Public institutions therefore require more regularized transition channels. Later sections develop this implication through interpretive pluralism, institutionalized rebalancing, public recognizability, and exit.

Threshold-like transitions provide one way to conceptualize these effects. Let q_t represent a composite condition relevant to transition, such as verified reciprocal experience, independent evidence, or protected interaction. A regime shift may occur when

$$q_t > q_c,$$

for some context-dependent threshold q_c . The notation should not imply that a single measurable scalar generally exists. Its purpose is to capture a practical point: many interventions may have little visible effect until enough mutually supporting conditions accumulate, after which relational change can become comparatively rapid. Governance evaluation that considers each intervention in isolation can therefore miss cumulative and nonlinear effects.

The reverse transition is equally important. A trusted public relationship can cross into distrust after a severe violation, a series of inconsistencies, or evidence that reorganizes the interpretation of earlier history. At that point the entire past may be re-read. Events once treated as benign can become evidence of a longer pattern. A basin transition can thus reorganize retrospective meaning as well as prospective expectation. This historical reinterpretation is one reason trust repair cannot be reduced to restoring a numerical value that existed before the violation.

4.5 Stability, Hysteresis, and Relational Lock-In

The language of hysteresis is useful when the path required to enter a relational regime differs from the path required to leave it. In a literal physical system, hysteresis has precise mathematical and empirical meanings. The present paper uses the term more cautiously to describe history-sensitive asymmetry: a relation at the same apparent level of current evidence can respond differently depending on the route by which that state was reached. Trust built through years of reciprocal experience may persist through disturbances that would prevent trust from forming in a new relation. Distrust generated by a serious integrity violation may also persist after some of the original conditions have been corrected.

Trust-repair research provides empirical reasons to take such history sensitivity seriously. Kim and colleagues found that responses to alleged violations had different repair effects depending on whether the violation concerned competence or integrity and on later evidence of guilt or innocence [20]. Gillespie and Dietz similarly argue that repair after organization-level failure is systemic and multilevel, involving organizational components and processes rather than a simple return of a dyadic attitude to its prior value

[13]. These findings do not establish hysteresis in a strict dynamical-systems sense. They support the more limited proposition that the route from trust to distrust and the route from distrust toward repaired trust can be structurally non-equivalent.

A useful schematic distinction is therefore

$$\Gamma_{T \rightarrow D} \neq \Gamma_{D \rightarrow T},$$

where Γ denotes a relational transition path. Repair may require acknowledgment, changed procedures, restitution, verification, renewed interaction, or time, even when the triggering violation was singular. Conversely, building trust initially may have required only limited low-risk cooperation. Different paths can also involve different actors: the institution that caused a failure may be unable to repair trust without independent oversight or third-party validation.

Relational lock-in occurs when stabilization is reinforced by structures that make alternative trajectories increasingly costly or inaccessible. Trust can create lock-in through dependence, concentrated delegation, proprietary infrastructure, closed information channels, or interpretive privilege. Distrust can create lock-in through segregation, defensive institutions, reputational exclusion, security dilemmas, or the disappearance of ordinary contact. In both cases, the stabilized relation can persist even after some of the conditions that originally justified it have weakened.

Lock-in should be distinguished from stability. Stability can be valuable. Durable trust can support coordination, reduce the need for constant verification, and make long-term cooperation possible. Durable caution can protect actors where risks are real. Lock-in becomes a governance concern when the cost of revision becomes detached from the quality of current evidence, or when the mechanisms necessary for reassessment are themselves disabled by the stabilized relation. The normative issue is therefore not persistence as such but persistence that erodes the system's capacity to re-recognize its own conditions.

This distinction suggests a preliminary criterion for later normative analysis. Let ρ_t represent the effective responsiveness of a relational regime to relevant new evidence and η_t the practical availability of alternative relational trajectories. A stable regime need not be problematic when both remain meaningfully positive. A self-sealing lock-in approaches the limiting condition

$$\rho_t \rightarrow 0, \quad \eta_t \rightarrow 0.$$

The expression is conceptual rather than operational. It marks the point at which stabilization ceases to function merely as continuity and begins to suppress the mechanisms of revision and escape.

The analysis of attractor-like dynamics therefore yields no general instruction to maximize trust, minimize distrust, or destabilize every durable relation. Its narrower contribution is to show why public trust has to be governed historically. Present attitudes are partly produced by earlier trajectories; those attitudes alter later evidence and interaction; and the resulting feedback can change the structure through which future revision becomes possible. Section 5 turns from this temporal stabilization to its interpretive consequences, examining how trust and distrust shape the reading of evidence and how interpretive authority can itself become self-reinforcing.

5 Interpretive Dynamics of Public Trust

The preceding section described trust and distrust as historically stabilized relational conditions. Stabilization, however, does more than influence whether actors cooperate, withdraw, delegate, or monitor. It also changes how later events are interpreted. The same statement, omission, error, or institutional response can acquire different significance when it is encountered within a relation already organized by trust or distrust. This section therefore turns from the persistence of relational trajectories to the interpretive mechanisms through which those trajectories are reproduced. The discussion proceeds from the effects of trust and distrust on evidential interpretation, to the emergence of interpretive privilege and authority, then to self-sealing structures, and finally to re-cognizability as a condition for continued revision.

Let an event or item of evidence at time t be denoted by e_t , and let $\mathcal{I}_t$ represent the interpretive mechanism through which its significance is constructed in a public relational setting. A minimal representation is

$$y_t = \mathcal{I}_t(e_t; T_t, D_t, H_t, A_t, C_t),$$

where y_t is the interpreted significance of the event, T_t and D_t are trust- and distrust-related conditions, H_t is accumulated relational history, A_t represents relevant authority structures, and C_t contains other contextual conditions. The distinction between e_t and y_t is essential. Public actors do not ordinarily respond to uninterpreted evidence. They respond to evidence as rendered meaningful through prior relations, categories, expectations, institutional roles, and credibility judgments. Trust and distrust are therefore not only outcomes of interpretation; once established, they become inputs into later interpretation.

5.1 Trust and the Interpretation of Evidence

Trust can widen the range of non-hostile interpretations available when conduct is ambiguous. A trusted colleague who responds late, an institution that makes a procedural error, or a public actor who gives an incomplete explanation may be interpreted against an accumulated history rather than judged from the isolated event. Such interpretive generosity can be socially productive. Long-term cooperation would be fragile if every local disturbance were treated as decisive evidence of unreliability. Trust can therefore function as a form of temporal contextualization: later evidence is situated within a history that gives actors reasons to distinguish anomaly from pattern.

This effect is consistent with accounts of trust that emphasize vulnerability, affective orientation, and relational expectations rather than mere probabilistic prediction [26, 19]. A trusting relation changes what an actor is prepared to risk and how the other's conduct is initially received. The interpretive consequence can be represented schematically by a trust-conditioned weighting function

$$w_t(e) = W(e \mid T_t, H_t, C_t),$$

where the evidential weight assigned to an event depends partly on the history and trust configuration in which it appears. The claim is not that evidence becomes arbitrary. It is that evidential significance in social relations is rarely independent of the relational conditions under which evidence is encountered.

Trust can also delegate interpretation. An actor may accept another's explanation of an event because the other has previously demonstrated competence, honesty, care, or epistemic access. In public life this can be indispensable. Citizens cannot independently reproduce every scientific assessment, administrative investigation, judicial finding, or technical judgment on which collective action depends. Some allocation of credibility and interpretive labor is therefore unavoidable. Trust can reduce the cognitive and institutional cost of repeatedly reopening every settled matter.

The same economy of interpretation creates a vulnerability. Once a source is trusted, its explanations can receive an initial credibility advantage over competing interpretations. A benign prior can become an interpretive privilege: adverse evidence is treated as exceptional, ambiguity is resolved charitably, and the trusted actor is granted greater latitude to explain its own conduct. Such latitude is not intrinsically defective. A mature relation often requires it. The problem arises when the privilege ceases to be defeasible—when trust no longer contextualizes evidence but determines in advance what the evidence is allowed to mean.

The distinction can be expressed by separating resilience from immunity. In a resilient trusting relation, sufficiently strong or repeated counterevidence can still alter the interpretation:

$$e_t^- \text{ of sufficient relevance} \quad \Rightarrow \quad \Delta T_{t+1} < 0 \text{ or } \Delta \mathcal{I}_{t+1} \neq 0.$$

In an interpretively immune relation, adverse evidence is systematically absorbed without a credible pathway of revision. The normative importance of this difference becomes clearer when trust is considered together with distrust.

5.2 Distrust and the Interpretation of Evidence

Distrust can perform an epistemically protective function. It can increase scrutiny, motivate independent verification, slow delegation, and prevent status or familiarity from being mistaken for reliability. In public governance, such caution is often institutionalized through auditing, opposition, appellate review, investigative journalism, adversarial procedure, and other practices that presume that claims should remain open to challenge. Distrust should therefore not be identified with irrationality, hostility, or epistemic failure.

Yet distrust also modifies interpretation. When suspicion is already established, ambiguous conduct can be assigned a more threatening meaning, while favorable conduct may be discounted as strategic performance. An apology can be read as manipulation, silence as concealment, self-defense as evasion, and cooperation as an attempt to gain access or legitimacy. The structural possibility is that distrust changes not only the prior probability assigned to an adverse hypothesis but also the standards by which later evidence is admitted as corrective.

Research on biased assimilation provides an adjacent cognitive mechanism. Lord, Ross, and Lepper found that people with strong prior positions evaluated congenial evidence more favorably and scrutinized disconfirming evidence more critically, with exposure to mixed evidence sometimes increasing polarization [24]. In political judgment, Taber and Lodge similarly reported prior-attitude, confirmation, and disconfirmation effects in the evaluation and selection of arguments [34]. These studies do not establish that public distrust is reducible to motivated reasoning. The present concern is relational and institutional as well as cognitive. They nevertheless illustrate a mechanism through which prior orientation can affect the treatment of subsequent evidence rather than merely its final conclusion.

A distrust-conditioned interpretation can therefore become asymmetric:

$$W(e^+ | D_t) < W(e^- | D_t),$$

where e^+ is potentially trust-restoring evidence and e^- trust-reducing evidence of comparable relevance. Such asymmetry may be justified when the costs of false reassurance are high or when past violations warrant a demanding evidential threshold. The governance problem begins when the threshold for revision becomes indeterminate or effectively unreachable. At that point distrust no longer operates as critical distance but as a structure that protects itself from correction.

Trust and distrust are thus symmetrical in one important respect. Either can preserve interpretive openness, and either can close it. Trust can protect a relation from overreaction while remaining responsive to evidence; distrust can protect a public from credulity while remaining responsive to evidence. Conversely, both can become mechanisms through which every new event is assimilated to an established narrative. The relevant normative axis is therefore not simply trust versus distrust, but revisable versus self-sealing interpretation.

5.3 Interpretive Privilege and Authority

Public interpretation is never distributed evenly. Expertise, institutional office, professional standing, network centrality, direct experience, rhetorical capacity, media access, and prior reputation all affect whose account receives attention and credibility. Some of these asymmetries are functional. A laboratory specialist may be better positioned than a lay observer to interpret a technical measurement; a direct witness may possess evidence unavailable to distant commentators; a court may be authorized to settle a legal dispute for practical purposes even though later review remains possible. Public governance cannot therefore treat every difference in interpretive weight as domination.

The difficulty is that credibility itself is relationally distributed. Fricker's account of testimonial injustice emphasizes that speakers can receive credibility deficits because of identity prejudice, while her broader analysis of epistemic injustice shows that power operates through practices of knowing and interpretation [10]. For the present argument, the important extension is that public trust can produce both deficits and surpluses of practical interpretive authority. A trusted institution or high-status actor may not merely be believed more often; it may acquire greater power to define what kind of event has occurred, which evidence counts as relevant, and which alternative explanations deserve serious consideration.

Let $a_i(t)$ denote the effective interpretive weight of actor i at time t . A public judgment may be represented schematically as depending on a weighted set of interpretations,

$$Y_t = \mathcal{G}(a_1(t)I_1(e_t), \dots, a_n(t)I_n(e_t)).$$

The notation does not imply that public judgment is literally an additive aggregation. It makes visible a basic asymmetry: equal formal opportunities to speak do not imply equal capacity to shape the resulting interpretation. Trust, status, and connection can increase $a_i(t)$, while distrust, stigma, exclusion, or weak access can reduce it.

Interpretive authority becomes especially consequential when it is reflexive. An actor with high credibility can influence the interpretation of events that bear on its own credibility. A respected organization may explain a failure in terms that preserve confidence; a distrusted target may be unable to offer an explanation without that explanation itself being treated as suspect. In both cases, the distribution of interpretive authority affects the future distribution of trust, and the future distribution of trust then affects interpretive authority:

$$A_t \longrightarrow \mathcal{I}_t \longrightarrow (T_{t+1}, D_{t+1}) \longrightarrow A_{t+1}.$$

This recursive relation is one route through which public trust can become a power structure rather than merely an attitude.

The appropriate response is not to abolish interpretive authority. Public life requires differentiated epistemic roles. The narrower requirement is that authority remain contestable, attributable, and revisable. An expert judgment can carry greater weight while still exposing its evidence and methods to qualified challenge; an institutional finding can organize action while retaining procedures for appeal; a witness can be believed without acquiring permanent immunity from contradiction. Section 6 develops the public distribution of such authority in greater detail. Here the concern is the point at which interpretive advantage becomes self-protective.

5.4 Self-Sealing Interpretive Structures

An interpretive structure becomes self-sealing when the range of possible observations that could meaningfully revise it approaches zero. In a strongly trusting configuration, evidence against the trusted actor may be reclassified as misunderstanding, exceptional error, hostile fabrication, or evidence of the accuser's unreliability. In a strongly distrusting configuration, favorable evidence may be reclassified as performance, manipulation, temporary compliance, or proof of greater sophistication. The content differs, but the structural form is similar: apparently heterogeneous inputs are mapped toward the same stabilized interpretation.

A schematic self-sealing condition can be expressed as

$$\mathcal{I}_t(e^+) \approx \mathcal{I}_t(e^-) \approx I^*$$

with respect to the direction of revision, where I^* is the dominant interpretation. This does not require the interpretations of the events to be verbally identical. It requires that neither class of event possess a practicable route to alter the dominant relational judgment. The mechanism can be cognitive, institutional, network-based, or some combination of the three.

Self-sealing structures are particularly likely when interpretive power and relational behavior reinforce one another. Consider a person who has acquired a negative reputation within an organization. Colleagues may reduce contact, making independent experience less available. The target's resulting anxiety or defensiveness may then be interpreted through the existing frame. Attempts at explanation can be discounted because the speaker already occupies a low-credibility position. The structure thereby produces new behavior and new evidence under conditions already shaped by the prior judgment. What appears to be repeated confirmation may partly reflect the system's own intervention in the relation.

The same structure can occur around trusted actors. Long-standing confidence can lead others to delegate investigation, interpretation, and communication back to the very institution whose conduct is at issue. If external evidence is then assessed primarily through that institution's categories, trust has begun to

govern its own conditions of revision. The problem is not simply excessive trust. It is a coupling between trust and interpretive authority that reduces the independence of corrective pathways.

For this reason, public governance requires a distinction between interpretive stability and interpretive closure. Stability means that every minor disturbance need not reopen the entire relation. Closure means that relevant disturbances no longer have a credible route into revision. Let ρ_t denote the effective responsiveness of an interpretive structure to relevant new evidence. A self-sealing configuration approaches

$$\rho_t \rightarrow 0,$$

not because no evidence exists, but because the mechanisms that assign evidence significance have become insensitive to it. The variable is conceptual rather than operational at this stage; Section 11 returns to the difficulties of measurement.

Self-sealing interpretation also helps explain why simply adding more information may fail to correct a public trust problem. If the interpretive mechanism remains unchanged, additional evidence can be assimilated through the same prior structure. Intervention may therefore have to alter who evaluates the evidence, how competing accounts are compared, what records remain accessible, or which procedures allow a settled interpretation to be reopened. In this sense, an interpretive problem can become a governance problem before it becomes a shortage-of-information problem.

5.5 Re-Cognizability and Interpretive Revision

The preceding analysis suggests that a healthy public relation cannot be defined by maximal trust, maximal skepticism, or permanent interpretive plurality. Public coordination requires provisional settlement. Decisions must sometimes be made, responsibilities allocated, and some interpretations treated as sufficiently reliable for action. The relevant question is whether such settlement preserves the possibility of later recognition: whether the relation can again be rendered intelligible as a historically generated structure and reconsidered in light of new evidence, new participants, or changed conditions.

Re-cognizability is used here to denote more than transparency. Transparency can make a present output visible while leaving the process that generated it difficult to reconstruct. A publicly re-cognizable judgment should retain, to a practicable degree, access to its evidential basis, the relevant sequence of interpretation, the identities or roles of significant interpreters, the procedures through which alternatives were excluded, and the conditions under which review can occur. The aim is not infinite archival completeness. Record preservation has costs, privacy constraints, and its own distributions of power. The narrower requirement is that consequential public interpretations should not become irreversibly detached from the histories through which they acquired authority.

A re-cognizable trust structure therefore permits actors to distinguish at least three questions that are often compressed into one: whether an interpretation is presently accepted, why it came to be accepted, and whether the conditions that once supported it continue to hold. These questions can diverge. A public may reasonably rely on an institutional judgment while also preserving the possibility that the judgment will later be revised. A long-standing distrust may remain prudent while the system continues to collect evidence that could eventually justify repair. Re-cognizability protects this temporal openness.

The concept also clarifies the paper's use of generativity. Interpretive openness does not mean that every interpretation is equally defensible or that meaning is unconstrained. A generative interpretive structure is structured yet unfinished: inherited evidence, concepts, procedures, and histories constrain what can plausibly be said, while future evidence and relations retain the capacity to modify the interpretation. In schematic form,

$$I_{t+1} = \Phi(I_t, e_{t+1}, H_{t+1}, C_{t+1}), \quad \frac{\partial I_{t+1}}{\partial e_{t+1}} \neq 0.$$

The derivative notation is heuristic. It marks the requirement that relevant future evidence retain some non-trivial capacity to alter interpretation.

This criterion applies symmetrically to trust and distrust. Good trust is capable of doubt; good distrust is capable of being changed by evidence. The normative opposite of a revisable trusting relation is therefore

not distrust, and the normative opposite of revisable distrust is not trust. In both cases the more serious danger is interpretive closure. Public governance should accordingly preserve channels through which settled judgments can be re-examined without requiring every participant to maintain continuous individual skepticism.

That final point is important for public institutions. A governance system cannot assume that every citizen, employee, official, or participant will possess exceptional diplomatic judgment, epistemic vigilance, or reflective capacity. Institutional design must carry part of the burden by preserving records, rights of reply, independent review, plural sources of evidence, and procedurally protected opportunities for revision. Such mechanisms do not eliminate power from interpretation. They provide conditions under which interpretive power is less able to become permanently self-sealing. Section 6 therefore moves from the internal dynamics of interpretation to the public distribution of interpretive power, examining how status, connection, reputation, and institutional position shape whose interpretations acquire effective authority.

6 Public Power and the Distribution of Interpretation

Section 5 treated trust and distrust as conditions that shape how evidence is interpreted. The present section shifts the unit of analysis from interpretive mechanism to interpretive distribution. Public interpretation does not occur among actors who possess equal access to evidence, equal credibility, equal network position, or equal capacity to make an interpretation circulate. Some actors can introduce categories that others subsequently inherit; some can make an interpretation visible across several communities; some can preserve a claim long enough for it to become reputationally consequential; and some can answer criticism from positions that already carry institutional or relational authority. The resulting asymmetries are not reducible to formal office. They may arise from status, expertise, connection, familiarity, audience reach, organizational role, accumulated trust, or combinations of these conditions.

The analytical objective is therefore limited but important. This section does not seek a general theory of public power, nor does it assume that interpretive asymmetry is inherently unjust. Public reasoning depends on differentiated epistemic roles, and some interpretations appropriately carry greater weight because their proponents possess relevant evidence, competence, responsibility, or direct experience. The governance problem begins when these differences become recursively self-reinforcing, when interpretive weight becomes difficult to revise, or when an actor's position allows it to shape the conditions under which its own credibility is assessed. The discussion proceeds through five connected dimensions: status and connection, publicity and participation, reputational reinforcement, relational exclusion and escape channels, and institutional mediation.

Let $V = \{1, \dots, n\}$ denote a set of public actors and $G_t = (V, E_t)$ their evolving relational network. For actor i , define an effective interpretive weight

$$w_i(t) = W(s_i(t), c_i(t), r_i(t), q_i(t), k_i(t), h_i(t), \dots),$$

where s_i denotes status, c_i network connection, r_i reputation, q_i institutional or procedural position, k_i relevant knowledge or evidential access, and h_i accumulated relational history. The notation does not imply that these heterogeneous quantities can be reduced to a single empirically measurable scalar. It serves only to mark that public interpretations are received through relationally distributed conditions. A public claim made by a highly connected and already trusted actor enters a different interpretive environment from an otherwise similar claim made by an isolated actor with little recognized standing.

6.1 Status, Connection, and Interpretive Weight

Status changes more than the probability that a statement will be heard. It can influence the initial presumption under which a statement is received, the amount of verification demanded before it is repeated, the seriousness accorded to competing accounts, and the extent to which ambiguity is resolved in the speaker's favor. An academic title, institutional office, professional reputation, organizational seniority, or recognized expertise may therefore operate as an interpretive multiplier. Such multipliers can be epistemically

useful. Public life would become impracticable if all claims had to be evaluated as though no actor possessed differential competence or prior reliability. The problem is not differential weight as such, but the conditions under which weight is assigned and revised.

Fricker’s account of testimonial injustice identifies one way in which socially distributed credibility can become unjust: prejudice may cause a speaker to receive less credibility than is epistemically warranted [10]. The present argument adds a relational concern that is not limited to identity prejudice. Interpretive weight may also be amplified through status and connection in ways that become recursively consequential. A high-status actor can make an interpretation more visible; greater visibility can produce more repetition; repetition can increase familiarity and perceived plausibility; and the resulting acceptance can further strengthen the actor’s reputation as an authoritative interpreter. Conversely, a low-status actor may face a credibility deficit before the substance of a claim is seriously considered.

Network position complicates this further. Centrality and power are related, but not in a mechanically monotonic way. Bonacich’s work on network centrality emphasizes that the relation between central position and power depends on the structure of the network and the dependence relations embedded in it [4]. This caution is important for public interpretation. A highly connected actor may enjoy broad reach, but a broker between otherwise disconnected communities may exercise a different kind of influence from an actor embedded in a dense cluster. Burt’s analysis of structural holes similarly shows how brokerage across disconnected groups can provide access to alternative information and confer social advantages on actors who span those gaps [7]. In interpretive terms, a broker may not simply possess more connections; it may be positioned to decide which interpretations cross community boundaries, how they are translated, and which alternatives remain locally invisible.

The effective weight of an interpretation therefore depends partly on topology. Let $m_{ij}(t)$ denote the capacity of actor i ’s interpretation to reach actor j at time t . Then public influence is not adequately represented by a speaker’s intrinsic credibility alone. A schematic expression is

$$\mathcal{W}_i(t) = \sum_j m_{ij}(t) w_i(t),$$

where $\mathcal{W}_i$ represents effective reach-weight rather than an empirical measure proposed for immediate application. Two actors with comparable local credibility can have very different public effects if one is connected to institutions, media, professional communities, or socially central intermediaries that permit its interpretation to travel farther and persist longer.

Connection also interacts with trust. If actor i is already trusted by several high-weight actors, then those relationships can function as secondary credibility channels. A claim does not need to persuade every member of a public independently; it may be accepted because trusted intermediaries accept or repeat it. In this way, interpretive authority can be inherited through a network. The structure is not necessarily illegitimate. Expert communities routinely rely on chains of specialized trust. Yet the same mechanism can make early reputational advantages difficult to dislodge, especially when later audiences encounter an interpretation only after it has passed through several trusted nodes.

This point matters for the analysis of public trust because trust itself can become part of the infrastructure through which interpretive weight is distributed. A trusted actor receives greater interpretive latitude; that latitude can help preserve its public standing; preserved standing then sustains future trust. The result is a coupled process

$$T_i(t) \longrightarrow w_i(t) \longrightarrow \mathcal{I}_t \longrightarrow r_i(t+1) \longrightarrow T_i(t+1),$$

where public trust, interpretive weight, and reputation are historically interdependent. The normative concern is not that such loops exist. Any stable public order contains feedback between recognized competence and subsequent credibility. The concern is whether the loop remains penetrable by relevant counterevidence and competing interpreters.

6.2 Publicity, Participation, and Interpretive Asymmetry

Publicity is often treated as a corrective to concentrated interpretation. Moving a dispute from a private hierarchy into a public forum can expose claims to broader scrutiny, allow additional evidence to enter, and

reduce the ability of one actor to control the informational environment. These are substantial advantages. Yet publicity should not be confused with equality of interpretive power. A process can be public while remaining highly asymmetric in whose claims are noticed, amplified, believed, remembered, or treated as agenda-setting.

Three dimensions should therefore be separated. *Publicity* concerns whether claims and procedures are exposed to a broader field of observation. *Participation* concerns whether relevant actors can enter that field and contribute claims, evidence, questions, or objections. *Interpretive distribution* concerns the effective weight and circulation of those contributions. The first two can improve without producing the third. A public hearing may permit all affected parties to speak while giving one institutional actor far greater prior credibility. A digital platform may permit universal posting while algorithmic, reputational, or network effects give some speakers orders of magnitude greater reach. A formally participatory organization may invite dissent while social dependence makes some forms of dissent costly enough that they remain practically unavailable.

This distinction is especially important when public governance is imagined as simply allowing “the public” to determine what is true. Public judgment is not generated by an undifferentiated mass. It is produced through sequences of attention, framing, repetition, contestation, and selective uptake. Even where no formal censor exists, the structure of visibility can determine which claims become candidates for collective assessment. Publicity therefore changes the field in which interpretation occurs, but it does not remove power from that field.

A minimal representation distinguishes presence from effective influence. Let $p_i(t) \in \{0, 1\}$ indicate whether actor i is formally permitted to participate, while $w_i(t)$ denotes effective interpretive weight. Formal inclusion gives

$$p_i(t) = 1,$$

but does not imply

$$w_i(t) = w_j(t)$$

for other participants j . Nor should equality of w_i always be the objective. A witness with direct evidence and an uninformed observer need not receive identical evidential weight. The narrower governance requirement is that the reasons for differential weight remain inspectable and that status or connection not become substitutes for evidential relevance.

Publicity can even intensify existing asymmetries. A high-status interpretation repeated in a public setting may acquire additional legitimacy from the appearance of collective acceptance. Actors who privately hold reservations may infer from the visible pattern of endorsement that their own information is exceptional or mistaken. Research on informational cascades provides a formal neighboring literature: Bikhchandani, Hirshleifer, and Welch show how individuals observing earlier actions can rationally follow those actions even when their private information points elsewhere, producing conformity that may be fragile with respect to new public information [3]. The present paper does not equate public trust dynamics with informational cascades, but the mechanism illustrates how publicly observable uptake can itself become evidence for later participants.

The same public field can therefore be both corrective and amplifying. Public exposure allows counterclaims to emerge, yet visible convergence can also discourage their expression or reduce their perceived credibility. This dual character supports a cautious governance principle: publicness is valuable because it creates opportunities for contestation, but those opportunities depend on the distribution of access, attention, and credibility inside the public field. The relevant object of governance is not merely whether a decision is public, but whether public processes preserve practically usable pathways for interpretation to be challenged.

6.3 Reputation and Distributed Reinforcement

Reputation connects past interpretation to future interpretive weight. Once an actor or institution acquires a reputation for competence, honesty, unreliability, danger, generosity, or misconduct, later observers rarely encounter subsequent conduct without that accumulated frame. Reputation therefore operates as a public

memory mechanism. It compresses distributed historical experience into a socially portable signal. Such compression can be useful: it allows actors to coordinate without independently reproducing every earlier interaction. At the same time, compression can detach a judgment from the evidence and contexts through which it was initially formed.

A simplified reputational update can be written as

$$r_i(t+1) = \mathcal{R}(r_i(t), e_i(t+1), \{m_{ji}(t)\}, \{w_j(t)\}, C_t),$$

where $e_i(t+1)$ denotes new conduct or evidence regarding actor i , m_{ji} the transmission of evaluations from other actors j , and w_j the interpretive weight of those evaluators. The important feature is that reputation is not updated from direct evidence alone. It is socially mediated. The testimony of a central, trusted, or institutionally authoritative actor can move reputation more strongly than the same assertion made by an isolated actor. Once the new reputation changes patterns of interaction, it can also alter the production of future evidence.

This creates the possibility of distributed reinforcement. Suppose an initial negative interpretation causes several actors to reduce contact with a target. Reduced contact lowers the production of direct corrective experience. Later entrants then observe an already existing pattern of distance and may treat that pattern as information about the target. Their own distancing becomes a further observable signal. No single participant needs to intend a coordinated exclusion for the structure to deepen. A parallel process can stabilize positive reputation: trusted endorsement produces access, access produces successful interactions, and successful interactions produce further endorsement.

The dynamic can be represented schematically as

$$r_t \longrightarrow E_t \longrightarrow H_{t+1} \longrightarrow r_{t+1},$$

where E_t denotes the pattern of available relational edges and H_{t+1} the history produced under that pattern. The direction of causality is therefore circular. Reputation influences which relations become possible, and those relations subsequently generate the evidence from which reputation is updated. A public reputation can thus become partially performative without being fictitious. Once it reorganizes access, attention, co-operation, and avoidance, it becomes a causal condition of the relation from which later judgments are drawn.

This mechanism helps explain why correcting a false or weakly grounded public judgment can be difficult even after its origin is challenged. Later observers may possess genuine evidence produced under the altered relational environment. A person who has been persistently excluded may become hesitant, defensive, or less informed about group norms; an institution subject to generalized distrust may become more secretive or legally defensive; a favored organization may gain resources that allow it to perform more competently. The later evidence is real, yet its production is historically entangled with the earlier interpretation. A governance analysis must therefore distinguish evidence about an actor from evidence partly generated by the relational conditions imposed on that actor.

Reputational reinforcement also clarifies why simple fact correction may be insufficient. If the network distribution of trust, attention, and contact remains unchanged, a correction can enter a field in which the original interpretation still possesses greater circulation and relational support. The problem is not solved by replacing public judgment with centralized adjudication. Rather, a mature public structure requires mechanisms through which reputation can be traced to evidence, revised over time, and prevented from permanently determining access to the very relations that could generate corrective information.

6.4 Relational Exclusion and Escape Channels

When an interpretation becomes dominant, exclusion can occur before any formal sanction is imposed. Others may reduce contact, withhold opportunities, stop sharing information, or avoid association. These changes alter the target's relational environment and can make exit from a distrust trajectory increasingly difficult. Section 4 described this as an attractor-like process; the present concern is topological. A subject

embedded in a network of mutually reinforcing distrust may have few relational paths through which an alternative interpretation can reach the wider public.

The problem is especially severe when every response is interpreted through the dominant frame. Silence can be treated as admission, protest as aggression, explanation as manipulation, and withdrawal as confirmation of deviance. Under these conditions the subject cannot easily produce a locally effective perturbation because the available actions are mapped back into the same interpretation. The absence of an escape channel is therefore not merely psychological isolation. It is a structural reduction in the number of routes through which new evidence or alternative interpretation can acquire public standing.

A relational escape channel can be defined provisionally as a connection, procedure, or intermediary that links a locally closed interpretive environment to a partially independent evidential or interpretive space. Let A be an actor located in a distrust-dominated subnetwork G_D . An edge to actor or institution B is potentially escape-enabling when B is not wholly dependent on the same interpretive closure and has access to another network G' , such that

$$A \leftrightarrow B \leftrightarrow G'.$$

The significance of B lies less in benevolence than in partial independence. A close friend who shares exactly the same excluded position may provide essential support but little public corrective reach. A high-status authority who merely substitutes its own unquestionable judgment may reverse the local outcome without improving the system's revisability. An effective bridge preserves an alternative path through which evidence can be heard, compared, and potentially returned to the original public field.

Network theory provides useful neighboring concepts. Granovetter's account of weak ties emphasizes their role in linking otherwise separated social regions and transmitting information across group boundaries [15]. Burt's work on structural holes likewise highlights the advantages held by actors positioned between disconnected groups [7]. The present notion of an escape channel is normatively narrower. It is concerned not with maximizing brokerage advantage but with preserving routes through which a locally stabilized interpretation can encounter relatively independent evidence and judgment.

This distinction matters in schools, workplaces, professional communities, and other bounded publics. It may be tempting to rely on an exceptional teacher, senior colleague, respected scholar, or socially central individual to intervene when a local distrust structure becomes self-reinforcing. Such intervention can be effective because an external or cross-cutting relation introduces evidence that the local field had discounted. Yet a just public order cannot make escape depend on the accidental possession of a powerful ally. Unequal access to high-status defenders would simply reproduce relational privilege at a second level.

The governance implication is therefore institutional rather than heroic. Systems should provide *institutionalized escape channels*: review outside the immediate chain of authority, routes for cross-unit evidence, rights of reply, protected access to records, procedures for correcting reputation, and opportunities for re-assessment by actors who are not fully embedded in the original interpretive structure. These mechanisms do not guarantee correct judgment. Their role is to prevent a local field from becoming the only field in which the subject can be known.

Escape channels also apply to excessive trust. A tightly trusting network can discount warnings from outsiders, rely on internal explanations, and interpret criticism as evidence of hostility. In such cases the relevant bridge may carry distrust or counterevidence into a closed trusting community. The symmetry is important: the purpose of an escape channel is not to replace distrust with trust or trust with distrust. It is to preserve the possibility that a stabilized relational interpretation can encounter a sufficiently independent perspective to become revisable.

6.5 Institutional Mediation of Interpretive Power

Public governance cannot eliminate interpretive power. Nor should it attempt to flatten all differences in credibility, expertise, responsibility, or social position. The practical task is to mediate the conditions under which interpretive weight accumulates and to preserve routes through which that weight can be challenged. This requires a shift from asking who should possess final interpretive authority to examining how interpretive roles are distributed, coupled, and reviewed.

One risk arises when several forms of authority become concentrated in the same actor or institutional location. Consider the combination

$$P_i = (p_i^{\text{agenda}}, p_i^{\text{evidence}}, p_i^{\text{interpretation}}, p_i^{\text{reputation}}, p_i^{\text{adjudication}}).$$

The components denote, respectively, capacities to define what issue is considered, control relevant evidence, provide the dominant interpretation, influence reputation, and settle the dispute. The notation is not intended as a cardinal power index. It identifies distinct functions that may become mutually reinforcing when tightly coupled. An organization in which one office can initiate an allegation, control records, define the meaning of those records, manage external communication, and make the final determination possesses a different closure risk from an organization in which these functions are partially separated.

Institutional mediation can therefore operate through differentiated roles. Independent review, adversarial presentation, appellate procedure, audit, ombuds functions, peer assessment, and protected response rights all create forms of structured friction. Their value is not that each participant is unbiased. Rather, they reduce the probability that one interpretive orientation simultaneously controls all stages through which a public judgment is generated. Cook, Hardin, and Levi's analysis of cooperation without trust is relevant at this point: institutional arrangements can support cooperation under conditions in which interpersonal trust is weak or unavailable [8]. Analogously, public interpretive governance should not depend on every participant possessing exceptional reflexivity. Procedures can externalize part of the burden of doubt, verification, and reconsideration.

The same logic explains why institutionalization should not be confused with centralization. A central reviewer can sometimes break a local interpretive closure, but the reviewer can itself become a new center of accumulated trust and authority. Public governance therefore requires reflexive mediation: mechanisms that correct one concentration of interpretive power should themselves remain reviewable. The objective is not an infinite regress of oversight bodies. It is a layered structure in which no single interpretive position is granted practical immunity merely because it occupies the role of corrector.

This approach also modifies the meaning of public balance. Balance need not mean that every actor possesses equal influence or that every interpretation remains permanently open. Collective action requires provisional decisions, and some interpretations must at times organize practice. A balanced public field is better understood as one in which asymmetries remain dynamically contestable: evidence can cross relational boundaries, reputations can be revised, excluded actors retain meaningful routes of response, and interpretive authorities can lose weight when the grounds supporting that weight deteriorate.

The resulting governance criterion can be stated as a condition on revision rather than equality. Let $w_i(t)$ be the effective interpretive weight of actor i . A healthy public structure does not require

$$w_i(t) = w_j(t) \quad \forall i, j.$$

It requires that $w_i(t)$ remain historically and evidentially responsive:

$$w_i(t+1) = \Omega(w_i(t), e_{t+1}, H_{t+1}, R_{t+1}, Q_{t+1}),$$

where new evidence, relational history, review, and institutional conditions retain a non-trivial capacity to alter interpretive weight. Persistent asymmetry may be legitimate; irreversible asymmetry is the more serious governance concern.

This formulation also clarifies why public institutions cannot assume the presence of a permanent central intervener capable of restoring balance whenever trust or distrust becomes pathological. Such an assumption merely relocates the problem of trust to the intervener. Public governance must instead distribute capacities for correction across procedures, networks, records, and review relationships. It should make rebalancing possible even when ordinary participants are neither diplomatic experts nor continuously reflective observers of the whole system.

The argument developed across Sections 5 and 6 can therefore be summarized in relational terms. Public trust and distrust condition interpretation; interpretation is distributed through unequal relations of status, connection, and reputation; those distributions alter future trust and distrust; and institutions

participate in the same dynamics when they mediate or concentrate interpretive authority. The next section introduces a preliminary field-governance representation for these coupled conditions. Its purpose is not to redescribe society as a literal physical field, but to provide a compact analytical language for interventions that act on distributed relational conditions rather than on isolated actors alone.

7 A Preliminary Field-Governance Representation

The preceding sections have treated public trust and distrust as relationally generated conditions whose effects are distributed through histories, institutions, reputations, and interpretive structures. This makes a purely actor-centred representation increasingly difficult. A public judgment may be altered by the status of a speaker, the density of a network, the availability of counterevidence, the design of a review procedure, the memory of prior cooperation, and the accumulated trust or distrust attached to institutions that mediate the dispute. No single variable is sufficient, and an intervention directed at one of these conditions may propagate through several others. The present section introduces a preliminary field-governance representation for this problem.

The representation is intentionally limited. It does not propose that public life is literally a physical field, that social variables possess a common metric, or that a complete action principle governs social evolution. The language of fields has a longer history in social inquiry, including Lewin’s attempt to describe behaviour through the configuration of a contemporaneous social and psychological field [23], while relational approaches such as Bourdieu and Wacquant’s methodological relationalism emphasize positions and relations rather than isolated substances [5]. Institutional analysis likewise treats action situations as embedded in rules, community attributes, material conditions, and feedback from outcomes [28]. The present use is narrower than each of these traditions. A *field* denotes a distributed relational condition whose local configuration affects the possible trajectories of public interaction. *Field governance* denotes intervention in those distributed conditions rather than direct command over every actor within them.

Five steps are developed. First, public trust is represented together with other relational fields over a common analytical domain. Second, an action-like functional is introduced as a compact description of coupling among those fields. Third, governance intervention is represented as a designed field that enters the same relational system it seeks to influence. Fourth, perturbative and response expansions are used to distinguish direct effects from higher-order consequences. Finally, public balance is treated as a scale- and observation-dependent property rather than a static equality of actors or fields. Each construction is provisional and is intended to organize later governance analysis rather than to provide an empirically complete model.

7.1 Relational Fields and Public Trust

Let $\mathcal{M}$ denote an analytical public domain. A point $z \in \mathcal{M}$ need not correspond to a physical location. It may index a situated relational position defined by time, institutional setting, communication channel, issue domain, and relevant actor configuration. The purpose of $\mathcal{M}$ is to prevent the analysis from assuming that trust is attached to actors alone. Public trust may differ across institutional sites, issues, audiences, and moments even when the nominal actors remain the same.

Over this domain, consider a collection of relational fields

$$\Phi(z) = (\phi_T(z), \phi_D(z), \phi_I(z), \phi_P(z), \phi_E(z), \phi_A(z), \dots),$$

where ϕ_T and ϕ_D denote trust- and distrust-supporting conditions, ϕ_I interpretive conditions, ϕ_P distributions of effective power, ϕ_E evidential accessibility, and ϕ_A attention or public visibility. The notation does not require each field to be a directly observable scalar. Some may ultimately be vectors, tensors, networks, probability distributions, or functionals over records and relations. Writing them as ϕ_i at this stage serves only to mark that their effects are distributed, locally variable, and mutually dependent.

The distinction between a field and an individual attitude is important. Suppose actor a reports a moderate level of trust in institution B . That report is one observation. The surrounding trust-supporting

field may additionally include prior institutional performance, shared procedures, familiar symbols, trusted intermediaries, available records, reputational histories, and the perceived cost of being wrong. Conversely, distrust may persist even where no individual actor actively disseminates a hostile interpretation because the relational environment continues to reproduce separation, selective exposure, institutional memory, or asymmetric access to corrective evidence. The field representation therefore treats attitudes as events within a broader relational configuration rather than as the whole object of analysis.

A local relational state can be written schematically as

$$X(z, t) = \mathcal{X}(\Phi(z, t), H(z, t), C(z, t)),$$

where H denotes relevant historical trajectories and C contextual conditions not yet represented explicitly as fields. This formulation preserves a central claim of Sections 3–6: current trust and distrust are historically generated, and their consequences depend on the configuration in which they occur. Identical values of ϕ_T need not produce identical outcomes when the histories, institutional constraints, or distributions of interpretive power differ.

The field language also permits coexistence. Trust and distrust need not be encoded as $\phi_D = -\phi_T$. A public may trust an institution’s technical competence while distrusting its motives, trust a procedure while distrusting its officeholder, or cooperate under safeguards while maintaining substantial suspicion. Accordingly,

$$\phi_T(z) > 0 \quad \text{and} \quad \phi_D(z) > 0$$

can hold simultaneously in different dimensions of the same relational configuration. This preserves the distinction developed earlier between trust, distrust, reliance, and institutionally protected cooperation.

The immediate advantage of the representation is therefore modest. It allows the unit of analysis to shift from “how much does actor a trust actor b ?” to “what distributed conditions make particular trusting, distrusting, interpretive, and cooperative trajectories more likely within this public domain?” The latter question is the entry point for field governance.

7.2 The Action and Coupled Relational Conditions

A field representation becomes useful only if the fields are allowed to interact. Let an action-like functional be written as

$$S[\Phi] = \int_{\mathcal{M}} \mathcal{L}(\Phi, \nabla\Phi, H, C) d\mu.$$

Here $\mathcal{L}$ is a relational density that summarizes local contributions and couplings, $\nabla\Phi$ represents variation across the analytical domain, and $d\mu$ is a measure appropriate to the chosen representation. The expression is not offered as a claim that public systems obey Hamilton’s principle. It is a bookkeeping device for stating that the trajectory associated with one relational condition cannot in general be evaluated independently of the others.

The density can be decomposed schematically as

$$\mathcal{L} = \sum_i \mathcal{L}_i[\phi_i] + \sum_{i,j} g_{ij} \mathcal{O}_i \mathcal{O}_j + \sum_{i,j,k} g_{ijk} \mathcal{O}_i \mathcal{O}_j \mathcal{O}_k + \dots,$$

where $\mathcal{O}_i$ denotes an observable or operator constructed from field ϕ_i and $g_{ij}, g_{ijk}, \dots$ denote context-dependent couplings. The expression should be read structurally rather than literally. Pairwise coupling can represent, for example, the interaction of trust with evidential accessibility. A higher-order term can represent a condition under which trust affects interpretation differently depending on public visibility and institutional power. The ellipsis leaves open still higher-order interactions rather than asserting that every mathematically possible interaction is empirically present.

Several mechanisms already developed in this paper can be expressed in this language. The loop

$$\phi_T \longrightarrow \phi_I \longrightarrow \phi_P \longrightarrow \phi_T$$

can represent a situation in which accumulated trust changes interpretive latitude, interpretive latitude changes effective authority, and increased authority subsequently reinforces trust. A corresponding distrust configuration may take the form

$$\phi_D \longrightarrow \text{relational withdrawal} \longrightarrow \phi_E \downarrow \longrightarrow \phi_D \uparrow,$$

where withdrawal reduces access to corrective evidence and thereby strengthens the conditions supporting distrust. These arrows are not substitutes for empirical identification. They indicate which couplings a concrete study would need to distinguish.

The action-like representation is useful partly because it resists single-cause governance. If an intervention changes ϕ_E by making evidence more visible, the resulting public effect depends on its couplings with attention, reputation, interpretive authority, and pre-existing trust. Greater visibility may improve review in one configuration and intensify reputational cascades in another. Similarly, increasing formal participation may have little effect when network and status couplings continue to concentrate effective interpretive weight. Governance therefore acts within a coupled system rather than upon a set of independent variables.

Historical dependence can be included by allowing the couplings themselves to evolve:

$$g_{i_1 \dots i_n}(t+1) = \Gamma_{i_1 \dots i_n}(g(t), H_{t+1}, \Phi_{t+1}).$$

The importance of this extension is conceptual. Repeated interaction does not merely move a public system through a fixed landscape; it can change which relations are salient, which institutions are trusted, which signals carry authority, and how strongly one condition affects another. The field configuration and the coupling structure are therefore both historically generated. This is the field-level analogue of the evolving transition rule introduced in Section 4.

7.3 Governance Intervention as a Designed Field

Field governance begins when a public intervention is represented not only as a command or decision but as a modification of distributed relational conditions. Let $\phi_e(z, t)$ denote a designed governance field. Depending on the application, ϕ_e may be instantiated through evidential disclosure rules, independent review, protected response rights, low-risk cooperative arrangements, procedural separation of roles, access to appeal, or other institutional designs. Calling it a field does not imply the presence of a single central designer. A governance field can be generated through multiple institutions, rules, infrastructures, and local practices whose combined operation changes the conditions under which public interaction occurs.

The total action-like functional becomes

$$S_{\text{tot}}[\Phi, \phi_e] = S_0[\Phi] + S_{\text{int}}[\Phi, \phi_e] + S_e[\phi_e],$$

where S_0 describes the existing relational configuration, S_e represents the institutional costs and internal dynamics of the intervention, and S_{int} represents its coupling with the fields already present. A minimal interaction term can be written as

$$S_{\text{int}} = \int_{\mathcal{M}} \left[\sum_i \lambda_{ei} \phi_e \mathcal{O}_i + \sum_{i,j} \lambda_{eij} \phi_e \mathcal{O}_i \mathcal{O}_j + \sum_{i,j,k} \lambda_{eijk} \phi_e \mathcal{O}_i \mathcal{O}_j \mathcal{O}_k + \dots \right] d\mu.$$

The higher-order terms are analytically important. A rule intended to improve evidential accessibility can alter trust differently when it is introduced through a highly trusted institution than when introduced through a distrusted one. Independent review can reduce local interpretive concentration while simultaneously increasing trust in the reviewing body, thereby creating a new centre of authority. A public reporting mechanism can improve accountability while also increasing strategic performance, reputational signalling, or selective attention. The intervention therefore couples to the conditions it seeks to govern and may generate effects not contained in its stated objective.

This formulation also prevents the governance field from occupying an imagined position outside the system. Once introduced, ϕ_e becomes part of the relational environment. Actors learn it, interpret it, rely

on it, contest it, route around it, or attempt to capture it. The intervention can consequently accumulate its own reputation and trust. Its later form should therefore be treated as endogenous:

$$\phi_e(t+1) = \mathcal{G}(\phi_e(t), \Phi(t), O(t), H(t)) ,$$

where $O(t)$ denotes observed outcomes. Governance is reflexive because the instrument changes the system and the changed system changes the practical meaning and operation of the instrument.

This point is compatible with institutional approaches in which action situations are structured by rules and contextual conditions while outcomes feed back into subsequent institutional arrangements [28]. The present representation differs in emphasis by foregrounding distributed relational conditions and their couplings. It asks not only which rule structures an action situation, but how that rule changes the surrounding fields of trust, distrust, evidence, attention, and interpretive authority through which later action situations are encountered.

A practical implication follows. The appropriate comparison is rarely between “intervention” and “no intervention” in the abstract. It is between alternative intervention fields embedded in different relational configurations. The same procedural design may stabilize public trust in one domain, produce defensive distrust in another, and become largely irrelevant in a third. Field governance therefore requires attention to where, when, and through which relations an intervention enters the system.

7.4 Higher-Order Coupling and Perturbative Effects

The distinction between intended and emergent effects can be represented through perturbation. Suppose $\bar{\Phi}$ denotes a background relational configuration and a governance intervention is introduced with small control parameter ϵ :

$$\Phi = \bar{\Phi} + \delta\Phi, \quad \phi_e = \epsilon\psi_e.$$

Around this configuration, an action-like expansion can be written as

$$S_{\text{tot}} = S^{(0)} + \epsilon S^{(1)} + \epsilon^2 S^{(2)} + \epsilon^3 S^{(3)} + \dots$$

The first-order term captures the most immediate response under the chosen representation. Higher orders capture effects that depend on interactions generated by the intervention itself. In governance terms, $S^{(1)}$ may correspond to the direct effect of an independent review procedure on evidential access, while $S^{(2)}$ may include changes in reputation, strategic disclosure, or trust in the reviewer that subsequently feed back into public interpretation. A still higher-order effect may appear when actors reorganize their networks or practices in anticipation of the procedure.

For an observable $Q_a(x)$, the same idea can be expressed as a response expansion:

$$\Delta Q_a(x) = \int \chi_a^{(1)}(x; y) \phi_e(y) d\mu_y + \frac{1}{2} \iint \chi_a^{(2)}(x; y, z) \phi_e(y) \phi_e(z) d\mu_y d\mu_z + \dots$$

The kernels $\chi^{(1)}, \chi^{(2)}, \dots$ represent first- and higher-order responsiveness. They need not be known analytically. Their conceptual role is to distinguish a local intervention from the distributed responses it induces. A non-zero $\chi_a^{(1)}(x; y)$ at distant or institutionally disconnected positions would indicate that the intervention propagates beyond its immediate site; a substantial $\chi^{(2)}$ would indicate that the effect depends on interactions among induced changes rather than on direct exposure alone.

This representation helps clarify why a governance measure that appears balanced at first order may become unbalanced after adaptation. A transparency rule can initially reduce information asymmetry, yet repeated adaptation may advantage actors with greater capacity to curate records, manage communication, or mobilize attention. A participation mechanism can initially broaden access, yet subsequent network effects can concentrate visibility around already central actors. A safeguard can enable cooperation under low trust, yet successful cooperation may later create dependencies that redistribute bargaining power. None of these outcomes follows necessarily. The point is that governance assessment should remain attentive to orders of effect beyond the immediately intended response.

The perturbative language has limits. Large interventions, abrupt crises, institutional collapse, or threshold transitions may not be well approximated by a series around $\bar{\Phi}$. Section 4 already emphasized that a system can cross basin boundaries and enter qualitatively different regimes. In such cases the background itself must be re-specified. Perturbation is therefore useful for studying local responses and emergent coupling, not for assuming that every governance transition is smooth or weak.

This limitation is substantively important. Public governance often operates precisely where small interventions accumulate until a threshold is crossed. A sequence of modest reputational cues may generate a distrust attractor; repeated low-risk cooperation may gradually make previously unavailable trust relations possible; incremental procedural concentration may eventually make external correction practically inaccessible. A field-governance representation should therefore combine local response analysis with sensitivity to regime change rather than treating one as a substitute for the other.

7.5 Scale, Observation, and Public Balance

The concept of balance becomes difficult once public trust is represented as a coupled field problem. Equal values of a selected variable at one level do not imply balanced relational capacity at another. Two actors may receive identical speaking time while differing greatly in network reach, reputational inheritance, evidential access, or institutional authority. Two groups may possess comparable formal rights while one has far greater ability to make its interpretation persist across time. Balance is therefore neither simple equality nor the disappearance of gradients.

Let $\mathcal{C}_\Lambda$ denote a coarse-graining operator at analytical scale Λ . The observed configuration is

$$\tilde{\Phi}_\Lambda = \mathcal{C}_\Lambda[\Phi].$$

At a fine scale, the analyst may observe local asymmetries among individual actors. At an organizational scale, those variations may average into apparently balanced departments. At a wider public scale, one institution may nevertheless dominate agenda setting or interpretive circulation. No single scale should therefore be presumed to reveal the complete governance condition.

Observation also requires a choice of representation. Let Γ denote an observational frame specifying which relational dimensions are made visible, how categories are defined, and which equivalences are treated as substantively meaningful. A balance observable can then be written as

$$\mathcal{B}_{\Lambda, \Gamma} = \mathcal{B}(\tilde{\Phi}_\Lambda; \Gamma).$$

The dependence on Γ is not an invitation to relativism. It is a reminder that apparently neutral measures can hide dimensions of power that were excluded from the observation. A procedure can appear balanced under a frame that records only formal participation and unbalanced under one that additionally observes agenda setting, evidence control, brokerage, or exit costs.

The language of an observational frame is intentionally weaker than a claim of literal gauge symmetry. A genuine gauge formulation would require specification of transformations under which physically or socially meaningful observables remain invariant. This paper does not establish such a symmetry. It uses a gauge-like intuition only to motivate a methodological requirement: governance claims should distinguish properties of the relational system from properties introduced by the analyst's coordinates, categories, or scale of observation.

A preliminary conception of public balance can now be stated. Balance does not require

$$\phi_i(z) = 0 \quad \text{or} \quad w_a = w_b$$

for all fields or actors. It refers instead to a configuration in which no relational field acquires persistent and practically unreviewable capacity to close the future evolution of the others. Trust may be high, authority may be asymmetric, and some interpretations may be provisionally dominant while the system remains balanced in this sense if counterevidence can still enter, interpretive weight can still change, affected actors retain meaningful routes of response, and governance mechanisms themselves remain reviewable.

This formulation connects field representation to the paper’s earlier concept of re-cognizability. A public configuration remains re-cognizable when later actors can reconstruct how a trust relation, reputational judgment, or interpretive authority was generated and can reopen it under materially relevant new conditions. Field balance is therefore not a terminal equilibrium. It is a property of an evolving system that preserves non-trivial capacities for revision, movement, and reconfiguration across relevant scales.

The resulting field-governance perspective is deliberately incomplete. It does not provide a universal metric of $\mathcal{B}$, an empirical procedure for measuring every ϕ_i , or a closed-form solution for S . Its contribution is narrower: public trust can be treated as one component of a distributed and historically coupled relational configuration; governance intervention enters that configuration as another field rather than as an external command; higher-order responses can make the effects of intervention diverge from first-order intentions; and claims of public balance must be indexed to scale and observation. Section 8 applies this representation to governance under conditions of trust and distrust, where cooperation, correction, and intervention must remain possible even when the relational field begins far from mutual confidence.

8 Governance under Trust and Distrust

The field-governance representation developed in Section 7 can now be applied to a narrower practical problem: how public cooperation, correction, and institutional learning remain possible when trust is low, unevenly distributed, or already organized into self-reinforcing patterns. The central premise of this section is that governance should not treat trust as a universally desirable stock that must simply be increased. Public systems frequently begin from conditions in which trust is unavailable, historically unreasonable, or normatively inappropriate. In such settings, the immediate task is often to make limited interaction possible while preserving the ability of participants to withhold trust, test claims, revise interpretations, and leave harmful relations.

This requires a distinction between five related governance functions. First, institutions can protect cooperation under low trust by limiting the losses associated with defection, error, or misinterpretation. Second, repeated protected interaction can generate relational histories from which trust may later emerge, although such emergence is neither automatic nor required. Third, where distrust has become attractor-like, corrective intervention may need to alter the evidential and relational conditions that reproduce it rather than merely exhort participants to think differently. Fourth, these corrective capacities should be institutionalized so that rebalancing does not depend upon the presence of an unusually trusted, connected, or diplomatically skilled intermediary. Finally, public judgments should remain re-cognizable: their evidential and interpretive histories should be sufficiently reconstructible that later actors can understand how they became authoritative and under what conditions they can be reopened.

The resulting orientation is therefore neither pro-trust nor pro-distrust in the abstract. It is concerned with the governance of trajectories. Trust can enable cooperation and later become dependency; distrust can protect against exploitation and later become self-sealing exclusion. Public governance must consequently preserve movement among relational states rather than presuppose a terminal condition in which one attitude permanently displaces the other.

8.1 Low-Trust Conditions and Protected Cooperation

A common difficulty in trust-centred governance is circularity. Cooperation is said to require trust, yet trust is expected to arise from successful cooperation. Where the initial relation is characterized by substantial uncertainty or suspicion, this produces a deadlock:

$$D_t \uparrow \longrightarrow C_t \downarrow \longrightarrow H_{t+1}^{\text{shared}} \downarrow \longrightarrow D_{t+1} \uparrow,$$

where D_t denotes distrust-supporting conditions, C_t cooperative interaction, and H_{t+1}^{shared} the stock of shared relational experience available for later interpretation. The problem is not merely attitudinal. Low cooperation reduces the opportunities through which claims can be tested, intentions can become legible, and new relational evidence can be produced.

A first governance response is therefore to separate *trust* from *cooperability*. Actors need not believe one another to be trustworthy in order to enter every form of joint action. Reversibility, verification, bounded commitments, collateral arrangements, third-party monitoring, staged implementation, and clear withdrawal procedures can reduce the amount of vulnerability that must be accepted at the beginning of a relation. Such mechanisms resemble what the earlier sections called protected reliance: cooperation is made possible partly because the relation has been structured so that failure need not impose an intolerable loss.

This point has a familiar analogue in international relations. Work on cooperation under anarchy has long examined how payoff structures, iteration, monitoring, and the relative costs of exploitation alter the feasibility of cooperation among actors who cannot rely upon a central guarantor [18, 29, 1]. Jervis's analysis of the security dilemma, for example, emphasizes that cooperation becomes more difficult when actors cannot distinguish defensive from offensive preparations and when the costs of exploitation are high [18]. The present argument does not import that framework wholesale into public governance. It takes from it a narrower design intuition: when trust cannot reasonably be assumed, governance can sometimes alter the relational consequences of being wrong.

Let p_t denote an actor's perceived probability that a counterpart will defect or that a cooperative arrangement will fail, and let L_t denote the loss that the actor expects to bear if this occurs. A simplified entry condition can be written as

$$p_t L_t \leq \theta_a,$$

where θ_a is the maximum risk actor a can presently accept. A trust-building strategy tries primarily to reduce p_t . A protected-cooperation strategy can instead reduce L_t , or distribute it through safeguards, while leaving p_t relatively high. The expression is not intended as a complete expected-utility model. Its analytical role is to show why an institution can make cooperation feasible even when the actor's underlying judgment of trustworthiness has barely changed.

This distinction matters normatively. Governance should not force an injured, historically subordinated, or otherwise vulnerable party to demonstrate reconciliation as the price of participation. A relation can be made safer while distrust remains intelligible. In diplomacy, organizations, public administration, and community conflict, the first attainable state may therefore be neither mutual trust nor shared identity, but a protected zone of limited cooperation in which the consequences of betrayal, error, and misrecognition are bounded.

The field representation makes this design problem more explicit. Let ϕ_e be an intervention field intended to enable cooperation under low trust. Its immediate purpose is not necessarily to increase ϕ_T . It may instead alter evidential access, reversibility, monitoring, or exposure:

$$\phi_e \longrightarrow (\phi_E \uparrow, L \downarrow, \text{reversibility} \uparrow) \longrightarrow C \uparrow.$$

Only later, and only under some relational histories, may this change the trust field. Protected cooperation is therefore not an inferior or incomplete form of trust. It is a distinct governance achievement that can preserve interaction where demanding trust would either be unrealistic or coercive.

8.2 Trust Formation through Relational Infrastructure

The fact that safeguards can substitute for trust at the point of entry does not mean that they are unrelated to later trust formation. Once protected cooperation becomes possible, actors acquire experiences that were unavailable under complete withdrawal. Repeated interaction can produce records of compliance, reveal competence, clarify expectations, and create opportunities for reciprocal adjustment. In Ostrom's behavioural account of collective action, reciprocity, reputation, and trust become mutually implicated in the development of sustained cooperation rather than appearing as isolated psychological variables [27]. The relevant implication here is historical: an institutional arrangement may first reduce the need for trust and subsequently help generate the evidence from which trust can become more reasonable.

A simplified sequence is

$$S_t \longrightarrow C_t \longrightarrow E_t^{\text{rel}} \longrightarrow H_{t+1} \longrightarrow T_{t+1},$$

where S_t denotes safeguards or relational infrastructure, E_t^{rel} the experience generated through cooperation, and H_{t+1} the updated relational history. The arrow from H_{t+1} to T_{t+1} should be read as a possibility rather than a law. The same history may support continued guarded cooperation, selective trust in one domain, or renewed distrust if the experience reveals opportunism. What infrastructure supplies is not trust itself but an environment in which trustworthiness can become more observable and relational interpretation can acquire new material.

This suggests a useful distinction between *trust cues*, *trust opportunities*, and *trust grounds*. Trust cues are symbolic, sensory, reputational, or institutional features that alter a person’s willingness to enter a relation. Section 3 discussed how status, setting, design, and even sensory conditions can affect such entry. Trust opportunities are low- or bounded-risk interactions through which another party’s conduct can be observed. Trust grounds are the historically accumulated reasons that later support a judgment that vulnerability is warranted. Governance becomes problematic when cues are engineered to mimic the affective experience of trust while opportunities for independent testing are weak and the grounds remain opaque.

Relational infrastructure can therefore be evaluated by the transitions it permits. A strong infrastructure should not merely make an institution appear trustworthy. It should facilitate interactions in which claims can be checked, commitments can be compared with conduct, errors can be acknowledged, and participants can revise their level of reliance. This is compatible with Hardin’s emphasis on trust as grounded in relationally situated expectations rather than generalized moral confidence [16], and with the distinction between political trust and governmental trustworthiness emphasized by Levi and Stoker [21]. The normative importance of infrastructure lies partly in preventing confidence from becoming detached from the conditions that would justify its continuation.

Procedural design can also affect whether an institution becomes a credible site for such learning. Research on procedural justice has shown that perceived fairness, respectful treatment, and legitimate procedure can matter for voluntary acceptance of legal authority rather than compliance being reducible to fear of sanction [35]. The present framework does not infer from this that fair procedure automatically produces trust. It instead treats procedural fairness as one possible relational condition that can alter how actors interpret institutional action and whether they remain willing to participate in the relation long enough for further evidence to accumulate.

Trust-generating infrastructure should therefore remain compatible with continued skepticism. A system that requires participants to stop checking once cooperation has begun converts a scaffold into a closure mechanism. A more robust trajectory is

protected entry $\longrightarrow$ repeated interaction $\longrightarrow$ historical evidence $\longrightarrow$ revisable trust or distrust.

The final term is intentionally disjunctive. Governance should make better-grounded relational judgment possible; it should not predetermine that the judgment must become more trusting.

8.3 Distrust Attractors and Corrective Intervention

The reverse problem arises when distrust has already become self-reinforcing. Section 4 described how distrust can reduce contact, narrow evidential access, change behaviour, and thereby generate further material that appears to confirm the initial interpretation. Section 5 added an interpretive mechanism: once distrust is sufficiently stabilized, favourable evidence may be discounted as strategic performance while unfavourable evidence is assimilated as confirmation. A distrust attractor is therefore not simply a large value of D_t . It is a relational regime in which ordinary local interactions increasingly reproduce the conditions that make distrust appear warranted.

In such a regime, asking the distrusted actor to “prove” trustworthiness can be structurally insufficient. Defensive explanation may be read as manipulation; silence may be read as admission; emotional response may be treated as evidence of instability; attempts to avoid a hostile environment may be interpreted as further suspicious conduct. The important point is not that these interpretations are always false. It is that a sufficiently closed interpretive field can make the evidential consequences of nearly every available response point in the same direction. The capacity for correction then depends less on the production of one additional statement and more on reopening the conditions under which statements are evaluated.

A corrective intervention should therefore be directed at the reproducing loop rather than at distrust as an attitude. Suppose a local distrust dynamic is schematically represented as

$$D_t \longrightarrow W_t \longrightarrow E_t^c \downarrow \longrightarrow I_t^D \longrightarrow D_{t+1},$$

where W_t is relational withdrawal, E_t^c access to corrective evidence, and I_t^D a distrust-weighted interpretation. Several intervention sites become available. An institution can reduce unnecessary withdrawal by creating safe, bounded contact; increase E_t^c through independent fact-finding or direct hearing; diversify interpretation by adding reviewers who are not embedded in the original relational cluster; or reduce the cost of leaving the relation while the dispute is unresolved.

In field terms, the intervention need not oppose ϕ_D directly. It can modify the couplings that make ϕ_D self-stabilizing:

$$\phi_e : \quad g_{D,W} \downarrow, \quad g_{W,E} \downarrow, \quad \phi_E^c \uparrow, \quad \rho_I \uparrow,$$

where ρ_I denotes responsiveness of interpretation to materially relevant new evidence. The notation is schematic, but it marks a substantive governance shift. The objective is not to manufacture a more favourable opinion. It is to restore pathways through which the opinion can become responsive again.

An external trusted actor can sometimes provide such a disturbance. A teacher, senior colleague, mediator, professional body, or person with a long independent history of interaction may possess evidence that the local group lacks and sufficient standing to prevent immediate dismissal of an alternative interpretation. Yet reliance on such actors creates its own inequality. Access to a highly connected protector is uneven, and the intervention may merely replace one interpretive authority with another. For public governance, the existence of an exceptional bridge node should therefore be treated as evidence of a missing institutional function rather than as a complete solution.

Corrective intervention also requires normative restraint. Distrust can be justified. A victim of exploitation need not be returned to contact with an aggressor merely because an analyst prefers relational openness. Likewise, an institution should not use mediation language to erase relevant power asymmetries or to compel reconciliation. The governance problem is narrower: where a public judgment affects rights, opportunities, reputation, participation, or access to common institutions, the conditions through which that judgment reproduces itself should remain open to proportionate evidential review. Restoring revisability is different from prescribing trust.

8.4 Institutionalized Rebalancing

If corrective movement depends upon the presence of an unusually wise mediator, trusted professor, respected elder, or diplomatically skilled public actor, then the system remains fragile. Public governance cannot reasonably assume that every participant will recognize an emerging attractor, understand the distribution of relational power, and know how to intervene at the correct moment. The cognitive and social burden is too high, and the very actors best positioned to rebalance a field may already be entangled in its trust and distrust structures.

A more durable approach is to externalize part of this reflexive burden into institutional design. Appeals, independent review, conflict-of-interest rules, separation of investigative and adjudicative functions, rights of reply, records of dissent, periodic reconsideration, and alternative reporting channels can each preserve a route that does not depend on the dominant local interpretation. Such mechanisms are not valuable because institutions are neutral by definition. Their value lies in creating additional relational paths through which concentrated interpretation can be tested.

This can be expressed as a network and field condition. Let $G_t = (V, E_t)$ denote the public relational network, and suppose a dominant cluster $K \subset V$ has accumulated high internal trust and convergent interpretation. An institutionalized rebalancing mechanism adds channels E_t^r that connect affected actors to review sites not wholly dependent on K :

$$E_{t+1} = E_t \cup E_t^r.$$

The importance of E_t^r is not that an external site must reverse the local judgment. It is that the local cluster no longer possesses exclusive control over the pathway by which the judgment can be reconsidered.

Hirschman's distinction among exit, voice, and loyalty offers a useful adjacent vocabulary for thinking about institutional responses to deterioration [17]. In the present framework, both exit and voice can be understood as relational escape capacities. Voice preserves a pathway for altering the relation from within; exit preserves a pathway for reducing exposure when internal correction fails. Neither capacity should be idealized. Exit may be prohibitively costly for vulnerable participants, while voice can be formally available but practically ineffective when interpretive authority is concentrated. Governance therefore concerns the real accessibility and consequence of these pathways rather than their formal declaration alone.

Institutionalized rebalancing does not imply equal interpretive weight for every claim. Evidence quality, expertise, proximity to events, conflicts of interest, and the reliability of prior records legitimately affect how claims are assessed. The relevant constraint is instead that interpretive weight remain revisable. Let $w_i(t)$ denote the effective public weight attached to interpreter i . A structurally open system permits

$$w_i(t+1) = \Omega(w_i(t), e_{t+1}, r_i, c_i, H_{t+1}),$$

where e_{t+1} denotes new evidence, r_i the interpreter's relevant reliability, and c_i conflicts or contextual conditions. A self-sealing hierarchy approaches a regime in which w_i changes little regardless of contrary evidence. Rebalancing preserves the possibility that authority can be strengthened, weakened, redistributed, or bypassed through publicly intelligible procedures.

The field perspective also cautions against coupling too many governance functions to one node. When the same actor can define the allegation, control evidential access, determine which interpretations are admissible, and issue the final judgment, multiple fields of power become mutually reinforcing. Separation of functions can therefore be interpreted as a low-order institutional approximation to field balance. It does not eliminate power, but it limits the probability that one local configuration becomes the sole route through which the public reality of a case is produced.

The deeper objective is not permanent equilibrium. Rebalancing is itself an intervention and can create new concentrations of authority. Review bodies can become insulated, procedural expertise can become exclusionary, and formal safeguards can be strategically used by actors with superior resources. Institutionalized rebalancing must therefore include review of the reviewing mechanism. In the notation of Section 7,

$$\phi_{e,t} \longrightarrow \Phi_{t+1} \longrightarrow \phi_{e,t+1},$$

so that the governance field remains part of the co-evolving system rather than an external source of final correction.

8.5 Public Re-Cognizability

The preceding mechanisms converge on a final governance requirement: public judgments should remain sufficiently re-cognizable that their histories can be reconstructed and materially relevant revisions can be made. Transparency is related to this requirement but does not exhaust it. A system can publish its final decision while leaving opaque the sequence by which evidence was selected, interpretations were privileged, dissent was excluded, and relational authority accumulated. Re-cognizability concerns the recoverability of that generative process.

Let a public judgment at time t be denoted J_t . A minimal provenance structure can be written as

$$\Pi(J_t) = \{e_k\}, \{I_k\}, \{a_k\}, \{r_k\}, H_t^J,$$

where $\{e_k\}$ is the materially relevant evidential record, $\{I_k\}$ the interpretations that substantially shaped the judgment, $\{a_k\}$ the actors or institutions exercising relevant authority, $\{r_k\}$ the rules or procedures through which the judgment was generated, and H_t^J its revision history. Public re-cognizability does not require every conversation or subjective impression to be permanently recorded. It requires enough structured provenance that a later reviewer can distinguish the judgment's present authority from the historical process that produced it.

This distinction is particularly important where trust itself has become evidence. Statements such as “this institution has always been reliable,” “everyone knows this person is difficult,” or “there is no reason to doubt the committee” compress prior relational histories into present interpretive weight. Such compression is unavoidable in complex systems; actors cannot reconstruct the full history of every public claim before acting. The governance problem arises when compression becomes irreversible. Re-cognizability preserves a route from the compressed public judgment back to a sufficiently rich account of its genesis.

Record design therefore becomes part of trust governance. Exhaustive recording is neither possible nor necessarily desirable. It imposes financial, cognitive, privacy, and administrative costs, and an indiscriminate archive can itself amplify power by privileging actors with greater capacity to produce documentation. A more proportionate approach is event-sensitive. Decisions that materially alter rights, access, reputation, institutional status, or the future evidential position of affected actors can trigger stronger provenance requirements than routine interactions. Disagreement, reversal, escalation, or reliance on exceptional authority can similarly trigger additional recording or review. The objective is to preserve future re-cognition at points where the system is most likely to become path dependent, while acknowledging that recording itself consumes resources and changes behaviour.

Publicity also should not be confused with truth by majority judgment. A public process does not become epistemically sound merely because many people can observe or vote on an interpretation. Publicity is valuable here because it can widen contestability, expose conflicts, and permit alternative evidence to enter. The truth of a factual claim remains dependent on the relevant evidential and inferential standards, not on the number of persons who endorse it. Public governance should therefore govern the conditions under which claims remain challengeable rather than attempt to make truth itself a plebiscitary product.

Re-cognizability can be summarized through three capacities. The first is *retrospective reconstruction*: later actors can identify how a judgment or trust relation acquired public force. The second is *responsive revision*: materially relevant new evidence can alter the judgment rather than being absorbed automatically into the established interpretation. The third is *prospective openness*: the system preserves routes through which future actors can reopen the relation when conditions change. Schematically,

$$\mathcal{R}_{extpublic} = R_{reconstruct} + R_{revise} + R_{reopen},$$

where the expression indicates dimensions of capacity rather than an additive empirical index.

This requirement brings the argument of the section back to its starting point. Low-trust governance should make cooperation possible without demanding premature trust; relational infrastructure should allow trust to emerge from inspectable histories rather than from cues alone; corrective intervention should reopen self-stabilizing distrust rather than manufacture favourable belief; and institutionalized rebalancing should provide public escape channels that do not depend on exceptional individuals. Re-cognizability links these functions across time. It allows the public system to stabilize enough to act while retaining the capacity to understand, revise, and sometimes leave the relational structures through which that stability was produced.

The next section turns from these governance capacities to the normative ambivalence of trust itself. The same trust that lowers coordination costs and supports collective action can also intensify dependency, facilitate collusion, externalize harm, and protect established interpretations from revision. A field-governance perspective therefore cannot end with successful trust formation. It must also ask what the stabilized trust relation enables, whose generativity it expands, whose alternatives it constrains, and whether the resulting configuration remains open to re-cognition and change.

9 Normative Ambivalence of Trust

The preceding sections have treated trust and distrust as relationally generated, historically stabilizing, and interpretively consequential conditions. This section turns to their normative evaluation. The central claim is deliberately limited: trust should not be treated as a good in itself, and distrust should not be treated as a harm in itself. Both alter what relations can do. Trust can lower the costs of coordination, support vulnerability, enlarge the range of joint action, and sustain forms of cooperation that would otherwise be difficult.

The same capacities can also deepen dependency, reduce monitoring, facilitate extraction, stabilize collusion, externalize harm to third parties, and protect established interpretations from revision. Distrust can interrupt these dynamics by preserving distance, demanding evidence, or motivating institutional safeguards, yet it can also harden into exclusion and self-sealing suspicion.

The normative problem is therefore not exhausted by the direction or intensity of trust. It concerns the trajectories and capacities that a trust or distrust relation makes possible, the distribution of vulnerability and authority within the relation, the effects imposed beyond it, and the degree to which the resulting configuration remains open to evidence and revision. This orientation is consistent with the vulnerability emphasized in philosophical accounts of trust [2] and with organizational work that explicitly identifies potentially harmful dynamics within trusting relations [32]. It also preserves the distinction developed earlier between trust and cooperation: complex public systems can support cooperation through institutional arrangements even where interpersonal trust remains limited [8].

Five dimensions are developed. The first concerns the capacity-expanding effects of trust. The second examines the transition from coordination to dependency and extraction. The third extends the unit of evaluation beyond the trusting dyad to collusion and externalized harm. The fourth develops the corresponding ambivalence of distrust. The final subsection proposes *revisable stability* and *generative openness* as provisional evaluative orientations. These are not offered as a complete moral theory of trust. Their purpose is narrower: to identify conditions under which stabilized trust or distrust remains compatible with continued public learning, contestation, and relational revision.

9.1 Trust, Cooperation, and Collective Capacity

Trust can enlarge the set of actions that a relation can sustain. Where actors expect one another to act within a tolerable range, they can disclose information, delegate tasks, accept temporary asymmetries, coordinate over longer horizons, and undertake projects whose success depends on conduct that cannot be continuously monitored. In this sense, trust is not merely an attitude toward another actor. Once enacted, it can become a relational capacity.

Let $K_{AB}(t)$ denote the effective joint capacity of actors or groups A and B at time t . A simplified relation can be written as

$$K_{AB}(t) = \mathcal{K}(T_{AB}(t), H_{AB}(t), I_{AB}(t), C_t),$$

where T_{AB} denotes trust-supporting conditions, H_{AB} their relational history, I_{AB} relevant institutional supports, and C_t the surrounding context. The expression does not imply that more trust monotonically produces more capacity. Its purpose is to mark that trust can reduce some forms of friction and thereby change the feasible action set of the relation.

This capacity-expanding role helps explain why trust is commonly valued. A relation in which every claim must be independently verified, every small commitment collateralized, and every action monitored can be prohibitively expensive. Trust allows part of this burden to be suspended. It therefore releases attention, time, and institutional resources for other forms of coordination. The benefits can be especially important where action is temporally extended or where participants possess different information and must rely on one another's competence and good faith.

Yet the same description already reveals the source of normative ambivalence. Trust increases what a relation can accomplish, but it does not determine the ends toward which that capacity is directed. Formally,

$$T_{AB} \uparrow \implies K_{AB} \uparrow$$

may hold under some conditions, while the normative sign of the resulting action remains unspecified. Increased joint capacity can support mutual aid, scientific cooperation, public administration, or peace-building. It can also support exclusion, coordinated deception, coercive enforcement, or organized violence. Trust is therefore better understood as a condition that can amplify relational agency than as a direct measure of the value of the resulting relation.

This distinction is particularly important for public governance. A policy that successfully increases trust between two institutions may improve coordination while simultaneously increasing their capacity to act without external scrutiny. A tightly integrated professional community may exchange information efficiently while becoming less responsive to criticism from outside the community. An alliance may reduce uncertainty among its members while increasing the vulnerability of actors beyond it. The normative evaluation of trust cannot therefore stop at the internal quality of the trusting relation.

The point also limits a common policy shorthand in which a decline of trust is treated as evidence of institutional failure and an increase of trust as evidence of improvement. Political trust may provide information about how citizens evaluate institutions, but trustworthiness and trust remain analytically distinct [21]. A public institution that becomes more trusted through improved conduct differs normatively from one that becomes more trusted through symbolic authority, selective disclosure, or successful control of interpretation. What matters is not only the level of trust attained but the relational process through which it is generated and the capacities that it subsequently enables.

9.2 Trust, Dependency, and Extraction

The capacity-enhancing effects of trust can alter the distribution of vulnerability within a relation. Trust commonly permits one party to reduce monitoring, disclose more information, delegate discretion, accept delayed reciprocation, or remain in a relation under uncertainty. These practices can be mutually beneficial. They can also create conditions in which the trusted party acquires increasing control over resources, interpretation, or exit.

A minimal dependence relation can be represented as

$$\Delta T_{A \rightarrow B} > 0 \implies \Delta V_A > 0, \quad \Delta M_A < 0,$$

where V_A denotes the vulnerability accepted by A and M_A the monitoring or verification burden retained by A . Neither change is inherently problematic. Vulnerability is partly constitutive of trust, and the reduction of monitoring is one of the mechanisms through which trust becomes practically valuable. The normative difficulty appears when the relation simultaneously raises the cost of exit or increases the trusted party's capacity to appropriate value generated through the relation.

Let E_A denote the effective ease of exit available to A and $P_{B|A}$ the effective relational power of B over decisions materially affecting A . A possible trajectory is

$$T_{A \rightarrow B} \uparrow \longrightarrow \text{delegation and disclosure} \uparrow \longrightarrow E_A \downarrow, \quad P_{B|A} \uparrow.$$

This is not an inevitable development of trust. It identifies a structural possibility: trust can change the relation that later determines how costly it is to withdraw trust. The history produced under trust can therefore become part of the mechanism that stabilizes dependency.

This dynamic is especially relevant where organizations or firms deliberately cultivate trust. Reliable service, familiar environments, institutional prestige, repeated successful interaction, and carefully designed symbolic or sensory conditions can lower the perceived risk of entering or deepening a relation. Such conditions may genuinely improve coordination. They may also generate what can be called *relationally accumulated trust*: a history in which the trusted organization becomes increasingly central to the actor's routines, information, identity, or access to alternatives. Once this history is established, the relation may support forms of value extraction that would have been rejected at the point of entry.

The temporal sequence matters. A relation can initially be mutually advantageous and later become extractive:

$$\text{trust formation} \longrightarrow \text{relational embedding} \longrightarrow \text{dependency} \longrightarrow \text{possible extraction}.$$

Research on embedded economic relations provides an adjacent structural caution. Uzzi finds that embedded ties can generate substantial cooperative benefits up to a point, while excessive embeddedness can expose firms to shocks and insulate them from information beyond the network [36]. Gargiulo and Benassi

similarly identify a trade-off between the safety afforded by cohesive networks and the flexibility required for adaptation [12]. These findings concern network embeddedness rather than trust as such, but they illustrate a general point relevant here: relational stabilization can produce benefits and rigidity through the same connective structure.

The normative response should not be to prohibit dependency. Care, education, finance, professional expertise, public administration, and intimate relations all contain forms of reliance that can become deep and asymmetrical. The more modest requirement is that trust-generated dependency remain visible and contestable. A relation becomes normatively troubling when the conditions that originally justified trust are no longer sufficient to revise it, when exit becomes prohibitively costly without corresponding safeguards, or when the trusted party can convert accumulated confidence into extraction while retaining the interpretive authority to describe the extraction as normal, necessary, or beneficial.

This links dependency to the interpretive analysis of Section 5. Trust can reduce suspicion in a way that supports ordinary cooperation, but it can also reduce the salience of warning signals. The literature on the “dark side” of trust similarly cautions against treating trust as inherently beneficial and emphasizes the obligations and vulnerabilities that can arise within trusting relations [32]. The field-governance implication is not to restore permanent suspicion. It is to preserve channels through which the relation can be re-evaluated before accumulated trust becomes a shield against the evidence needed to assess its continued justification.

9.3 Trust, Collusion, and Externalized Harm

A second limitation of trust-centred evaluation appears when the unit of analysis is expanded beyond the trusting parties. A relation can be internally cooperative and externally harmful. If A and B develop high mutual trust, they may coordinate more effectively, share information more freely, and accept greater interdependence. From the perspective of the dyad, these are familiar indicators of successful trust formation. From the perspective of a third party C , the same increase in joint capacity can impose new risks.

The distinction can be represented by separating internal relational effects from external effects:

$$\Delta K_{AB} > 0, \quad \Delta \mathcal{G}_A \geq 0, \quad \Delta \mathcal{G}_B \geq 0,$$

while allowing

$$\Delta \mathcal{G}_C < 0,$$

where $\mathcal{G}_i$ denotes, provisionally, the generative opportunities or relational capacities available to actor or group i . No aggregation rule is assumed. The notation simply prevents the internal success of A – B coordination from being treated as sufficient evidence of public benefit.

This problem appears in many domains. Two firms may trust one another sufficiently to coordinate practices that disadvantage competitors or consumers. Organizational insiders may develop strong loyalty while jointly suppressing information harmful to the organization. Political or military partners may reduce uncertainty within an alliance while increasing the ability to coerce or attack actors outside it. Informal networks may provide members with reliable mutual support while excluding those who lack the relevant connections. In each case, the trust relation is real and may even possess many virtues internally. The normative problem arises because the strengthened relation changes the conditions of action for those who are not parties to it.

The term *collusive trust* can be used provisionally for trust that materially facilitates coordinated conduct whose costs are transferred to actors outside the trusting relation. The adjective is not intended to imply illegality. A jointly harmful relation can remain formally lawful, and a legally prohibited relation can contain genuine loyalty and mutual care among its participants. The purpose of the distinction is to prevent an internal relational good from exhausting the public evaluation.

This requires a change in the boundary of observation. If the analytical frame contains only A and B , their growing trust may appear as a monotonic improvement. Once C and the surrounding institutional field are included, the evaluation can change. In field terms, the local deepening of ϕ_T between two nodes may be accompanied by increased gradients in power, evidential access, or exposure elsewhere. Public trust

governance should therefore resist a dyadic metric in which the quality of the internal relation is treated as the whole normative object.

The same point applies to interpretive communities. High internal trust can support rapid collective sense-making, but it can also make external criticism easier to discount. If insiders possess strong confidence in one another's motives and competence, evidence originating outside the group may be assigned systematically lower weight. This does not make internal trust wrongful. It shows how a relation that initially supports epistemic cooperation can also stabilize a boundary around legitimate interpretation. The relevant governance concern is therefore not simply whether the group trusts itself, but whether the trust relation remains permeable to evidence and to the claims of those affected by the group's decisions.

A field-governance perspective consequently evaluates at least three relational surfaces: the internal distribution of vulnerability within the trusting relation, the external effects imposed beyond it, and the interpretive boundary through which affected outsiders can contest its actions. This broader frame is essential if trust is to be evaluated as a public rather than merely interpersonal phenomenon.

9.4 Distrust, Critical Distance, and Closure

The normative ambivalence of trust has a corresponding implication for distrust. Distrust can be destructive, humiliating, and socially isolating. Sections 4–6 described how an initial distrust condition can become self-reinforcing through withdrawal, selective evidence, reputation, and interpretive closure. In organizational or public settings, these dynamics can make an actor's ordinary responses appear confirmatory: silence can be read as admission, defence as evasion, and emotional response as further evidence of unreliability. Such a configuration narrows the routes through which distrust can be corrected.

It would nevertheless be mistaken to conclude that distrust should simply be minimized. Distrust can preserve critical distance. It can motivate verification, auditing, separation of powers, contestation of authority, demands for reasons, and safeguards against irreversible delegation. Cook, Hardin, and Levi explicitly argue that cooperation in complex societies often depends on institutional mechanisms that reduce the need for interpersonal trust and that distrust can motivate the creation of reliable constraints [8]. Lewicki, McAllister, and Bies likewise reject a simple unidimensional continuum in which trust and distrust are merely opposites and emphasize the possibility of their coexistence within complex relationships [22].

The distinction needed here is therefore between *critical distrust* and *self-sealing distrust*. Critical distrust withholds confidence while remaining responsive to reasons. It can be represented schematically by

$$D_t > 0, \quad \frac{\partial I_{t+1}}{\partial e_{t+1}} \neq 0,$$

where I_{t+1} is the subsequent interpretation and e_{t+1} materially relevant new evidence. Distrust remains present, but evidence can alter the interpretation. Self-sealing distrust instead approaches a regime in which positive, neutral, and negative events are all assimilated into the same hostile frame. The practical problem is not the existence of suspicion but the disappearance of conditions under which suspicion could become unwarranted.

This distinction mirrors the earlier analysis of trust. Trust can become self-sealing when contrary evidence is automatically discounted because the trusted actor is presumed incapable of the alleged conduct. Distrust can become self-sealing when favourable evidence is discounted as strategic performance, concealment, or manipulation. In both cases the relation loses interpretive responsiveness. The normative axis therefore does not run simply from distrust to trust. It also runs from revisability to closure.

A two-dimensional representation makes this clearer. Let s denote the orientation from distrust to trust, while r denotes interpretive revisability. Then relations with similar values of s can differ sharply in r . High trust with high revisability can support cooperation while preserving scrutiny; high trust with low revisability can become dogmatic or dependent. High distrust with high revisability can sustain critical oversight; high distrust with low revisability can stabilize exclusion. The analytical purpose of the second dimension is to prevent the normative evaluation from being collapsed into the amount of trust alone.

This also clarifies why public institutions often require both trust and institutionalized doubt. Courts, scientific communities, regulatory systems, and administrative organizations cannot function if every participant treats every statement as presumptively fraudulent. Yet their credibility also depends on review, appeal, replication, audit, conflict-of-interest rules, and other practices that refuse to make trust immune from examination. Such arrangements are not failures to achieve complete trust. They are mechanisms through which trust and doubt can coexist without either becoming terminal.

9.5 Revisable Stability and Generative Openness

The preceding analysis suggests that the normative objective of public trust governance should not be maximal trust, maximal distrust, or permanent relational fluidity. Public systems require stabilization. Actors need to rely on procedures, institutions, collaborators, and accumulated judgments without reconstructing their entire evidential history at every moment. Trust is valuable partly because it makes such compression possible. A system in which every relation is continuously reopened would impose unsustainable cognitive and institutional costs.

The problem is therefore to distinguish stability that supports continued action from stability that eliminates the conditions of its own revision. This paper uses *revisable stability* to describe a configuration that is sufficiently stable to support cooperation yet remains responsive to materially relevant evidence, alternative interpretation, and changed circumstances. A provisional representation is

$$S_R = f(S, \rho_E, X, V, R_c),$$

where S denotes relational stability, ρ_E responsiveness to relevant evidence, X meaningful possibilities of exit or institutional redirection, V access to voice and contestation, and R_c the capacity for re-cognition of the relation's history. The expression is not an empirical index. It identifies dimensions that can be examined separately rather than treating stability itself as evidence of legitimacy.

This concept can be extended through *generative openness*. A relation is generatively open when its stabilization does not pre-empt all later trajectories. It contains enough structure to coordinate action while preserving the possibility that new evidence, new participants, changed conditions, or previously excluded interpretations can alter the relation. In this sense, openness is not equivalent to indeterminacy. A completely unstructured relation may be incapable of sustaining cooperation at all. Generativity requires constraints, but the constraints should not permanently monopolize the future.

For trust, generative openness means that confidence can coexist with the possibility of doubt. For distrust, it means that critical distance can coexist with the possibility of restored confidence. The desirable structure can therefore be stated compactly as

$$\text{trust capable of doubt} \quad \text{and} \quad \text{doubt capable of trust}.$$

The phrase should not be read as a demand for symmetrical movement at every moment. Some harms warrant durable distrust, and some institutions may justifiably acquire stable confidence. The requirement is procedural and relational: the system should preserve intelligible conditions under which either orientation can be reconsidered when materially relevant circumstances change.

This is also where the normative evaluation of third-party effects re-enters. A trust relation may remain revisable internally while continuing to externalize substantial harm. Generative openness must therefore be assessed across the relevant relational boundary. A relation that expands the options of its members by systematically eliminating the options of outsiders cannot be evaluated solely by its internal flexibility. Conversely, governance that protects outsiders by making all deep cooperative relations impossible would sacrifice important forms of collective agency. The field perspective is useful precisely because it allows these effects to be considered as coupled rather than independently optimized conditions.

The resulting orientation is intentionally modest. It does not provide a complete ordering of competing values, nor does it identify a universal threshold at which trust becomes excessive. It proposes instead a set of diagnostic questions for later governance analysis: whether trust or distrust remains evidentially revisable; whether vulnerability and exit costs have become asymmetrically concentrated; whether collective capacity

is being used to externalize harm; whether interpretive authority has become insulated from challenge; and whether the relational structure still permits affected actors to re-enter the process by which its legitimacy is assessed.

Under this view, the normative achievement is not a final state called “trust.” It is a public relational configuration capable of acting, stabilizing, learning, and reopening itself. Trust and distrust are both resources within that configuration and both potential sources of closure. Their value depends on the generative relations they sustain and on whether those relations preserve the capacity to become otherwise.

10 Implications for Public Governance

The preceding sections suggest a shift in the object of public trust governance. If trust and distrust are historically generated, mutually reinforcing, interpretively consequential, and embedded in wider distributions of power, governance cannot be reduced to raising or lowering a scalar quantity called trust. Nor can it assume that the relevant problem is always located inside individual attitudes. The practical object is broader: the relational conditions under which actors enter cooperation, interpret evidence, accumulate or lose authority, become locked into established trajectories, and retain or lose meaningful routes of revision and exit.

This shift does not imply that public institutions can or should design every relation. It instead identifies a narrower governance responsibility. Institutions can shape background conditions that make some trajectories easier to enter, some forms of concentration easier to challenge, and some stabilized judgments easier to reopen. In the field representation introduced in Section 7, this means that governance acts less like a command that directly sets the value of trust than like a perturbation of conditions under which several coupled relational processes evolve. Because those interventions are themselves interpreted, strategically used, and historically sedimented, their effects cannot be assumed to remain identical to their initial purpose.

The implications developed here are therefore diagnostic rather than comprehensive. Five areas are considered. The first concerns the governance of initial relational conditions, including low-risk entry into cooperation. The second concerns interpretive concentration and the distribution of effective authority in public judgment. The third addresses relational lock-in and the conditions under which stabilization becomes difficult to revise. The fourth concerns revision, exit, and the preservation of institutional escape channels. The final subsection brings these elements together as field-level conditions for public trust. The objective is not to specify an optimal level of trust, but to identify public arrangements that allow trust and distrust to remain usable without becoming self-sealing.

10.1 Governance of Initial Relational Conditions

Initial relational conditions matter because the first interactions in a public or institutional relationship can influence the history that later participants treat as evidence. Section 3 distinguished trust cues from trust grounds, while Section 4 described how small differences can be amplified through feedback and path dependence. The governance implication is not that institutions should manufacture a favorable first impression. It is that they should pay attention to the structure of entry into relations, especially where the cost of a failed first interaction is high or where early interpretations are likely to be widely reproduced.

Let x_0 denote the initial relational state of an interaction and let $\Gamma(x_0)$ denote the set of trajectories that are practically accessible from it under prevailing institutional conditions. A governance intervention need not select a final trajectory. It may instead alter the accessible set,

$$\Gamma(x_0) \longrightarrow \Gamma_{\phi_e}(x_0),$$

where ϕ_e represents a designed intervention or institutional condition. The practical aim can be modest: increase the availability of trajectories in which actors can test cooperation, obtain interpretable evidence, and withdraw from early failure without catastrophic loss.

This is especially important under low trust. Where actors face substantial uncertainty about one another, demanding trust as a precondition for participation can freeze a relation before any corrective history is generated. Institutional safeguards can instead reduce the cost of being wrong. Limited commitments, reversible steps, third-party verification, clear procedural boundaries, and staged forms of cooperation can permit interaction even where confidence remains weak. This extends the argument developed in Section 8: cooperation can precede trust, and repeated cooperation can subsequently become evidence relevant to trust. Institutional analysis has long emphasized that rules and monitoring arrangements affect what forms of cooperation are feasible under uncertainty [28, 27]; the present framework treats these arrangements additionally as conditions that can influence the later relational history of trust and distrust.

The design problem is therefore not simply to maximize comfort or perceived safety. A highly reassuring environment can encourage participation while simultaneously making later withdrawal difficult or obscuring relevant risk. Initial trust cues can also be distributed unequally. Institutional prestige, professional titles, architectural settings, technical language, ceremonial procedure, or established reputation may place some actors inside a favorable interpretive frame before their conduct has been observed. Such cues may be useful shorthand, but they can also cause symbolic authority to substitute for evidence. Public governance should therefore distinguish between conditions that reduce unnecessary friction and conditions that suppress the visibility of uncertainty.

A useful initial-condition design can be described by several properties. First, the downside of exploratory participation should be bounded where possible. Second, early conduct should generate evidence that participants can later inspect rather than merely impressions that are difficult to contest. Third, the structure should preserve meaningful alternatives so that a participant does not have to deepen dependency merely to evaluate the relation. Fourth, the terms of interaction should be intelligible enough that a later failure can be attributed with some precision rather than absorbed into generalized suspicion. Fifth, asymmetries in vulnerability should be visible and, where they are substantial, institutionally compensated.

These properties also help distinguish trust-supporting governance from manipulative trust production. Manipulation seeks to alter the subject's disposition while restricting access to the conditions under which that disposition could be reconsidered. Trust-supporting governance, by contrast, can make cooperative entry easier while preserving evidence, alternatives, and revision. The difference is therefore not whether an environment influences perception; every public environment does. The difference lies partly in whether the influence preserves the subject's later capacity to understand and revise the relation.

Procedural legitimacy matters at this stage because actors often evaluate institutions not only by substantive outcomes but by the fairness and intelligibility of the procedures through which decisions are made [35]. A fair procedure does not guarantee trust, and trust should not be inferred merely from compliance. Yet procedures that make reasons, roles, and review paths legible can reduce the need for actors to rely on opaque personal assurances. They can also make later trust more evidentially grounded because the history of the relation contains interpretable institutional conduct rather than only symbolic cues.

Initial relational governance should therefore be understood as the design of entry conditions rather than the manufacture of confidence. Its task is to make low-risk learning possible, to prevent first impressions from becoming unnecessarily irreversible, and to allow participants to accumulate a history that can support either warranted trust or warranted distrust. The direction of the eventual judgment remains open.

10.2 Governance of Interpretive Concentration

Public judgment is never formed under conditions of equal interpretive influence. Section 6 described how status, connection, brokerage, reputation, and institutional position affect the weight carried by an interpretation. Publicity and participation can widen access to a discussion while leaving these asymmetries substantially intact. The governance problem is therefore not solved by allowing more actors to speak. It concerns the distribution of effective capacity to define the issue, specify relevant evidence, frame ambiguous events, and stabilize a public interpretation.

Let $w_i(t)$ denote the effective interpretive weight of actor i at time t within a defined public domain. The value is not assumed to be directly measurable. It represents a composite of visibility, credibility, network

access, institutional position, agenda-setting capacity, and the ability to translate claims into consequential decisions. A schematic concentration measure can be written as

$$\mathcal{C}_I(t) = \sum_{i=1}^N \tilde{w}_i(t)^2, \quad \sum_{i=1}^N \tilde{w}_i(t) = 1.$$

The expression is not proposed as an empirical index of public legitimacy. Its purpose is to make one structural point explicit: the number of participants can increase while effective interpretive concentration remains high.

Interpretive asymmetry is not in itself evidence of injustice. Expertise, direct experience, formal responsibility, and access to relevant evidence can justify unequal weight. A public institution would be dysfunctional if every statement had to be treated as epistemically identical regardless of competence or evidential position. The more precise concern is whether high interpretive weight becomes insulated from the conditions that originally justified it. Expertise can become authority, authority can shape which evidence is collected, and the resulting evidential environment can then reinforce the authority that helped produce it.

This possibility is especially important when several dimensions of power are coupled in the same actor or institution. A person who can initiate an allegation, control access to records, frame the relevant categories, determine who is heard, and participate in the final adjudication possesses more than a high level of credibility. Several stages of interpretation have become structurally dependent on the same node. The risk is not merely bias in one decision; it is the loss of independent routes through which bias could become visible.

Public governance can respond by differentiating functions rather than by trying to equalize all voices. Separation among evidence collection, interpretation, review, and final adjudication can create friction against self-confirmation. Requirements to explain and justify conduct before a forum are central to relational accounts of accountability; Bovens characterizes public accountability in terms of a relation in which an actor must provide information and explanation, a forum can question and judge, and consequences may follow [6]. The present analysis adds that the forum itself should not be treated as automatically neutral. Its position in the wider interpretive field, its access to evidence, and its dependence on existing authority also require examination.

Participation similarly needs more structure than an invitation to speak. Fung emphasizes that participatory arrangements differ in who participates, how participants communicate and make decisions, and how deliberation is connected to public action [11]. These dimensions are relevant here because interpretive power can concentrate at any of them. A forum may be broadly open yet permit only a narrow set of claims to affect institutional decisions. It may include affected actors while requiring them to translate their experience into categories controlled by specialists. It may collect testimony while leaving agenda-setting and evidential standards unchanged.

The governance objective is therefore better described as *contestable asymmetry*. Actors may possess unequal interpretive weight, but the sources of that weight should remain identifiable and challengeable. Counterevidence should not become inadmissible merely because it conflicts with a trusted authority. Affected actors should have routes to place alternative interpretations into a forum capable of producing consequences. Brokerage across otherwise separated groups can also matter because network position affects exposure to alternative information and perspectives [7, 15]. This does not imply that diversity automatically corrects error. It means that a system with multiple evidential and relational pathways is less dependent on the continued reliability of a single interpretive center.

This approach also avoids requiring every participant to become an expert in public balancing. A well-designed public field should not depend on each citizen, employee, student, or service user constantly recalculating the distribution of interpretive power. Institutional arrangements can externalize part of that burden through independent review, reason-giving requirements, differentiated roles, conflict-of-interest rules, and access to alternative forums. The relevant principle is not that all interpretations should receive equal authority, but that no temporary interpretive advantage should become equivalent to immunity from revision.

10.3 Governance of Relational Lock-In

Stability is necessary for public life. Institutions cannot function if every judgment, delegation, partnership, or rule must be reconstructed from the beginning at each moment. Sections 4 and 9 nevertheless showed that stabilization can become lock-in when the history generated by a relation reduces the system's later capacity to respond to evidence, alternative interpretation, or changed circumstances. The governance challenge is therefore to distinguish useful stabilization from closure.

Relational lock-in can arise through several mechanisms. Repeated cooperation can deepen dependence on a trusted institution. Repeated suspicion can reduce direct contact and thereby reduce corrective evidence. Reputation can cause new participants to inherit an established judgment before interacting with the subject of that judgment. Technical or procedural investments can raise switching costs. Interpretive authority can accumulate around actors who become increasingly central to the production and validation of evidence. These mechanisms can reinforce one another even when no participant intends the overall result.

A provisional representation can write the degree of lock-in at time t as

$$\mathcal{L}_t = L(H_t, E_t, C_t, 1 - \rho_t, K_t),$$

where H_t denotes accumulated historical dependence, E_t the effective cost of exit or redirection, C_t the concentration of relevant relational pathways, ρ_t interpretive responsiveness to materially relevant evidence, and K_t coupling among mutually reinforcing mechanisms. The expression does not imply that the components are commensurable. It identifies distinct sources through which a relationship can become difficult to alter.

Path dependence is especially relevant because self-reinforcing processes can make initially contingent arrangements increasingly costly to reverse [30]. In trust governance, however, reversal should not be treated as an intrinsic good. A durable public institution may deserve stable confidence precisely because it has accumulated a long history of reliable conduct. The governance problem appears when the same history that supports confidence also prevents the institution from being re-evaluated after its conduct changes.

This distinction suggests a role for what can be called *anti-lock-in frictions*. Periodic review, sunset provisions, rotation of evaluators, independent audit, interoperable systems, alternative service channels, and preservation of cross-group contact can prevent a historical arrangement from becoming the only feasible arrangement. Such mechanisms impose costs. Constant review can undermine continuity, consume attention, and create strategic uncertainty. The relevant design problem is therefore not to destabilize relations continuously but to preserve plausible routes of reopening at moments when the expected cost of continued closure becomes substantial.

Intervention thresholds may accordingly depend on events rather than fixed schedules alone. A stable relation may need little active disturbance until a material change occurs: a new conflict of interest, a substantial expansion of authority, a serious adverse event, a change in the evidential basis of a judgment, or a sudden increase in exit costs. Event-sensitive review can reduce the burden of constant reopening while preserving a structured response to circumstances that materially alter the relation.

Lock-in governance also requires attention to the direction of causality. A highly trusted organization may face little scrutiny because it is trusted, but low scrutiny can itself reduce the production of information needed to evaluate whether trust remains warranted. Conversely, a distrusted actor may be heavily monitored, and the monitoring environment may generate more observable minor deviations, thereby sustaining distrust. Governance should therefore examine not only the current judgment but the information-production regime generated by that judgment.

This point is important for public balance. A relation can appear stable and consensual while the conditions for discovering disagreement have been progressively removed. Conversely, a relation can appear conflictual because it preserves active channels for complaint and revision. Observable harmony is therefore a poor substitute for examining whether the field retains meaningful corrective capacity. The relevant goal is not a low level of visible dispute but a structure in which stabilization does not eliminate the routes by which materially relevant dispute can become consequential.

10.4 Governance of Revision and Exit

Revision requires more than the formal possibility of changing one's mind. In public systems, judgments are often embedded in records, reputations, organizational roles, access rights, and networks of subsequent decisions. Once these downstream effects accumulate, a later correction can leave the earlier relational consequences largely intact. A governance system that permits revision only at the level of statement therefore may remain practically irreversible.

The distinction between voice and exit is useful here. Hirschman's classic analysis treats exit and voice as different responses to organizational decline [17]. For trust governance, neither is sufficient alone. Exit protects actors from being indefinitely bound to a relation, but exit can be prohibitively costly or can remove precisely those actors whose experience is needed for institutional learning. Voice allows problems to be raised within a relation, but voice without an independent route of consequence can become merely symbolic. Public governance should therefore preserve both the capacity to contest and the capacity to redirect or leave when contestation fails.

This requirement is particularly important where a person becomes the object of a stabilized public or organizational interpretation. An appeal process that merely returns the case to the same interpretive network may not constitute a meaningful escape channel. Similarly, an independent reviewer who has no access to the relevant records or no authority to alter downstream consequences may provide procedural appearance without practical revisability. A relational escape channel requires at least some independence in evidence access, interpretive authority, and remedial capacity.

Revision also depends on memory. Section 8 introduced public re-cognizability as a stronger condition than transparency: later actors should be able to reconstruct how a judgment was generated, which evidence was available, which alternatives were considered, and which authority authorized the result. This does not require recording everything. Comprehensive recording can be financially expensive, cognitively burdensome, invasive of privacy, and itself a source of power. The governance problem is therefore to determine when preservation of provenance is valuable enough to justify its costs.

An event-sensitive recording rule can be represented schematically as

$$\mathcal{R}(e) = 1 \quad \text{when} \quad \mathbb{E}[B_{\text{future revision}} \mid e] > C_{\text{record}}(e),$$

where $\mathcal{R}(e)$ indicates that event e triggers enhanced record preservation, $B_{\text{future revision}}$ denotes the expected value of retaining sufficient provenance for later review, and C_{record} includes financial, administrative, privacy, and informational costs. The formula is not intended as a literal cost-benefit algorithm. It expresses a design principle: recording obligations can increase when decisions materially alter rights, reputation, institutional status, access to future opportunities, or the evidential position from which later review would occur.

Such records should preserve distinctions among allegation, evidence, interpretation, decision, and later revision. When these categories collapse into a single institutional memory, an early interpretation can survive long after the evidence that originally supported it has been weakened or rejected. A corrected judgment should therefore be capable of propagating through the same systems that propagated the original judgment. Revision without downstream repair risks producing a formally corrected but relationally unchanged outcome.

The same principle applies to positive trust. Institutions that accumulate extensive authority because of a history of reliable conduct should retain channels through which new evidence can alter their status. Trustworthiness should not be transformed into a permanent exemption from accountability. Conversely, distrust should not permanently disable an actor's ability to produce evidence of changed conduct. A revisable public order requires both directions of movement to remain conceptually and procedurally available, even though particular cases may justify durable confidence or durable caution.

This is also where epistemic injustice becomes institutionally relevant. Fricker's account emphasizes ways in which credibility and interpretive resources can be unfairly distributed [10]. The present framework does not reduce public trust problems to epistemic injustice, but it shares a concern with whether affected actors can enter the processes through which their conduct and experience are interpreted. Revision procedures should therefore examine not only whether a forum exists but whether the actor has

intelligible access to the language, evidence, and institutional routes through which the forum can actually reconsider the relation.

Governance of revision and exit is thus not a residual safeguard added after trust has been established. It is part of the conditions under which trust can remain legitimate over time. A relation that cannot be reconsidered when its evidential basis changes has ceased to function merely as trust and has begun to acquire the structure of closure.

10.5 Field-Level Conditions for Public Trust

The preceding implications can be summarized as conditions on a public relational field rather than as instructions to maximize trust. A public system contains multiple overlapping processes: cooperation, reputation, evidence production, interpretation, authority, vulnerability, exit, and institutional memory. Trust and distrust emerge within these processes and subsequently modify them. Governance therefore concerns the quality of the field through which these feedbacks occur.

Let

$$\mathcal{F}_t = (A_t, E_t, I_t, P_t, X_t, R_t, M_t)$$

represent a provisional set of field-level conditions, where A_t denotes accessibility of low-risk relational entry, E_t accessibility and provenance of relevant evidence, I_t plurality and contestability of interpretation, P_t the distribution and coupling of effective interpretive power, X_t meaningful exit and redirection possibilities, R_t revisability and re-cognizability, and M_t institutional memory sufficient to support later reconstruction. These variables are neither exhaustive nor proposed as a universal measurement model. They provide a compact representation of the governance dimensions developed across Sections 3–10.

Under this representation, public trust is better treated as an emergent relational outcome than as a direct policy target. A well-governed field may sometimes produce greater trust because institutions repeatedly demonstrate competence, fairness, and responsiveness. It may sometimes produce durable distrust because evidence warrants caution. It may also sustain cooperation while leaving interpersonal or political trust limited. The normative achievement lies in the quality of the process through which these orientations are generated and revised, not in their convergence toward a predetermined value.

This perspective also changes the meaning of balance. Public balance is not a static equality of power, voice, or trust. Nor is it the disappearance of asymmetry. It is closer to a condition in which no locally stabilized configuration acquires unrestricted capacity to close the routes through which other materially relevant information, interpretations, or affected actors can enter. Because public fields differ across scale and observational frame, the diagnosis of balance must remain explicit about what domain, time horizon, and relational boundary are being considered.

The governance field ϕ_e introduced in Section 7 should accordingly be treated as reflexive. An intervention that creates an independent review body, transparency requirement, participatory forum, or verification mechanism adds a new actor and new relations to the field. Over time the intervention can accumulate authority, develop routines, become dependent on particular information sources, or itself become resistant to revision. Public governance therefore requires governance of the governance mechanism. The relevant question is not only whether an intervention corrects the current imbalance, but whether it preserves conditions under which its own authority can later be reconsidered.

This reflexivity helps explain why permanent reliance on exceptional individuals is an unstable solution. A respected teacher, senior official, public intellectual, or well-connected intermediary can sometimes interrupt a harmful distrust attractor by introducing credible counterevidence or reopening a closed interpretation. Such intervention can be valuable, but a just public order should not make access to correction depend on the accidental presence of a powerful protector. The structural task is to institutionalize enough alternative pathways that ordinary participants retain some route to review even when no unusually influential ally is available.

The resulting governance orientation can be stated in restrained terms. Public institutions should seek to preserve conditions for protected entry, evidential learning, contestable interpretive asymmetry, bounded lock-in, meaningful voice and exit, event-sensitive provenance, and reflexive review. These conditions do

not eliminate power, distrust, error, or conflict. They reduce the degree to which any one historically produced relation can become immune to the processes through which public reality is continually reconstructed.

A compact normative expression is therefore not

$$\max T,$$

but rather the maintenance of a field in which

$$\rho_E > 0, \quad X > 0, \quad V > 0, \quad R_c > 0,$$

for materially affected relations, where ρ_E denotes responsiveness to relevant evidence, X meaningful exit or redirection, V effective voice and contestation, and R_c re-cognizability. These inequalities are conceptual rather than quantitative thresholds. Their purpose is to state that a public relation should retain some non-zero capacity to respond, redirect, contest, and reconstruct its own history.

This is a deliberately limited conclusion. It does not specify how competing claims to evidence should ultimately be adjudicated, how much asymmetry is legitimate, or how the costs of preserving revision should be distributed. Those questions require additional political, legal, ethical, and empirical analysis. The contribution of the field-governance perspective is narrower: public trust should be governed through the conditions of its emergence and revision rather than treated as a quantity to be maximized. Trust can support public cooperation precisely because it stabilizes relations; governance becomes necessary because the same stabilizing capacity can also close them.

11 Limitations and Research Boundaries

The preceding sections have used relational, dynamical, and field-oriented language to reorganize several problems that are often treated separately: the formation of trust and distrust, their historical stabilization, their effects on interpretation, the unequal distribution of interpretive authority, and the design of public conditions under which judgments remain revisable. This vocabulary is useful only if its boundaries are kept explicit. The present paper does not offer a general field theory of public life, a complete dynamical model of trust, or a universal normative criterion for public governance. Its formal expressions are preliminary representations intended to clarify dependencies and possible feedbacks that would otherwise remain implicit.

Four groups of limitations are especially important. The first concerns the field representation itself and the risk of turning a relational heuristic into an unwarranted social ontology. The second concerns the use of attractor language in open, historically changing systems. The third concerns measurement and empirical identification: most of the variables introduced in earlier sections are latent, relational, scale-dependent, and difficult to isolate causally. The fourth concerns normative evaluation. Re-cognizability, revisability, plurality, and exit are valuable governance considerations, but none is sufficient by itself to determine justice, legitimacy, or the appropriate substantive outcome of a public dispute. These limitations do not render the framework unusable. They specify what would have to be established before stronger theoretical or empirical claims could be made.

11.1 Limits of the Field Representation

The field representation introduced in Section 7 should first be understood as an analytical compression. It provides a way to describe conditions that are distributed across persons, institutions, symbols, records, networks, and public environments rather than located in a single actor. To write

$$\Phi(x, t) = (\phi_T, \phi_D, \phi_I, \phi_P, \phi_E, \dots)$$

is therefore not to claim that trust, distrust, interpretation, power, or evidence exist as physical fields in the same ontological sense as fields in fundamental physics. The notation says less: values relevant to public

trust may vary across relational and institutional locations, those values may be mutually dependent, and an intervention at one location may have consequences elsewhere.

This distinction matters because a compact field notation can hide the heterogeneity of what it represents. A change in ϕ_P , for example, might correspond in one case to a formal change in legal authority, in another to increased network centrality, in another to reputational advantage, and in another to control over a technical platform. These mechanisms need not be commensurable. Treating them as components of one representational object is useful for asking whether they interact, but it does not establish that they share a common unit, measurement scale, or causal mechanism. The same caution applies to trust and distrust. Sections 2–5 treated them as relational orientations with domain-specific and interpretive consequences; a field notation should not silently convert them into homogeneous quantities.

The action-like functional

$$S[\Phi] = \int_{\mathcal{M}} \mathcal{L}(\Phi, \nabla\Phi, H, C) d\mu$$

requires similar restraint. Its role in this paper is organizational. It makes it possible to state that multiple relational conditions may contribute jointly to a trajectory and that a governance intervention may interact with several of them at once. The paper does not derive an Euler–Lagrange equation for public trust, identify a conserved quantity, or claim that public systems optimize or extremize S . Social processes contain strategic adaptation, reflexive interpretation, institutional rule changes, discontinuities, and historical events that may alter the effective state space itself. A literal variational reading would therefore require assumptions that have not been supplied here.

The representation also depends on a chosen boundary. Any field Φ is defined over some domain $\mathcal{M}$, yet the relevant public domain is rarely self-evident. A workplace dispute can depend on professional associations outside the workplace; a national trust relation can depend on transnational media, financial institutions, diasporic communities, and historical memories; an online controversy can cross several platforms and legal jurisdictions. A boundary that is analytically convenient may omit precisely the relation through which power, evidence, or interpretive authority enters the system. Field analysis therefore requires an explicit statement of spatial, institutional, temporal, and relational scope rather than assuming that the chosen domain is naturally complete.

A second boundary problem concerns the observer. The variables included in Φ , the scale at which they are aggregated, and the distinctions drawn among evidence, interpretation, reputation, and authority are themselves analytical choices. Section 7 used a gauge-like analogy to emphasize that an apparent balance can depend on the observational frame. That analogy should remain limited. The paper has not specified a transformation group under which social descriptions are formally equivalent, nor has it derived gauge-invariant social observables. The more modest claim is that public balance should be tested under multiple plausible descriptions and scales rather than inferred from one convenient coordinate system.

The designed intervention field ϕ_e also creates a conceptual danger if it is imagined as external to the relations it governs. Public institutions, review bodies, transparency requirements, verification procedures, and participatory forums are themselves social relations. They accumulate reputation, develop routines, depend on information sources, and can become objects of trust or distrust. Their introduction changes the field partly because they become new elements of it. The notation

$$\Phi_t \longrightarrow \Phi_{t+1}(\phi_{e,t})$$

therefore has to be supplemented by the reverse dependence

$$\phi_{e,t+1} = G(\phi_{e,t}, \Phi_{t+1}, H_{t+1}),$$

where the intervention itself is historically modified by the system in which it operates. This reflexivity is consistent with institutional approaches that treat rules, action situations, and feedback as mutually conditioning rather than externally fixed [28]. It also limits any interpretation of field governance as technocratic control from outside the governed domain.

Finally, the field vocabulary can overstate continuity. Public life contains abrupt legal decisions, organizational dissolution, technological discontinuities, war, scandal, death, migration, and other events that

can change relations faster than a smooth field approximation suggests. Continuous notation can remain useful locally while event-based representations are more appropriate elsewhere. A mature version of the framework would likely require both: fields for distributed conditions and events for discontinuous transformations. The present paper does not yet provide that combined formalism.

11.2 Limits of Attractor Language

The attractor language developed in Section 4 is likewise provisional. In a strict dynamical-systems sense, an attractor is defined relative to a specified state space and evolution rule. Public trust rarely satisfies the corresponding stability assumptions. Actors learn, institutions change, relevant variables appear or disappear, and the function governing transition can itself evolve with history. The paper therefore speaks of *attractor-like* stabilization rather than asserting that a measured public-trust system possesses mathematically demonstrated attractors.

This qualification becomes especially important when the transition rule is historical:

$$x_{t+1} = F_t(x_t, u_t, C_t), \quad F_{t+1} = \Psi(F_t, H_{t+1}).$$

If F_t changes substantially over time, a basin identified at one period need not persist. A relation can appear strongly stabilized because the observation window is short, because countervailing relations are temporarily inactive, or because the institution has not yet encountered a sufficiently large perturbation. What looks like an attractor may therefore be metastability, institutional inertia, repeated strategic equilibrium, reputational lock-in, or simple persistence under a narrow range of conditions. These possibilities should not be collapsed merely because they produce similar trajectories over a finite interval.

Path dependence also does not by itself establish attraction. A historical sequence can constrain later options without generating convergence toward a stable region. Pierson’s account of increasing returns in politics is useful precisely because it emphasizes mechanisms by which early events can alter later costs and alternatives [30]; this is related to, but not identical with, an attractor claim. In empirical work, the mechanisms producing persistence should therefore be identified before assigning dynamical language to the observed pattern.

A similar caution applies to basin boundaries. Section 4 used the idea of a basin to describe regions from which trust-supporting or distrust-supporting trajectories become easier to reproduce. In actual public relations, the relevant boundary may be fuzzy, multi-dimensional, and dependent on actors’ interpretations of the same event. There may be no single threshold T_c or D_c . Different relational domains can also coexist: an institution may be trusted for technical competence, distrusted for distributive fairness, relied upon because alternatives are absent, and accepted procedurally despite low affective confidence. A one-dimensional phase portrait would erase these distinctions.

The source of perturbation should not be oversimplified either. Earlier sections emphasized that deeply stabilized distrust may require an external relational channel, independent review, or a new source of evidence to reopen interpretation. This does not imply that escape is always externally caused. Actors can reinterpret their own histories, internal coalitions can change, latent contradictions can become salient, and routine processes can generate unexpected evidence. Conversely, an ostensibly external intervener can already be embedded in the same network of assumptions and incentives. The distinction between endogenous and exogenous perturbation is therefore itself relational.

Hysteresis is another useful but limited analogy. Trust repair research shows that the effect of a violation and the efficacy of apology, denial, or institutional repair can depend on the perceived basis of trust and on the type of failure [20, 13]. Such findings support history-sensitive and asymmetric dynamics in some settings, but they do not establish a universal law that trust always takes longer to build than to destroy or that the return path necessarily differs from the entry path. The present framework should therefore treat hysteresis as an empirical possibility to be tested within a specified relation, not as a defining property of trust.

The strongest use of attractor language in this paper is consequently diagnostic. It directs attention to self-reinforcing feedback, path dependence, reduced sensitivity to ordinary perturbation, and the possibility

that interpretations help reproduce the conditions under which they remain persuasive. These properties can be investigated without presupposing a literal attractor. Future work would have to specify state variables, temporal resolution, transition rules, and criteria for convergence before stronger dynamical claims could be justified.

11.3 Measurement and Empirical Identification

The framework presently faces a substantial measurement problem. Trust, distrust, interpretive authority, re-cognizability, and relational lock-in are not directly observable in the form in which they appear in the conceptual model. Observable indicators such as survey responses, compliance, transaction frequency, communication patterns, network position, or willingness to accept vulnerability capture only parts of these constructs. The distinction between trust and reliance is especially important: behavior that appears trusting may result from lack of alternatives, coercion, contractual safeguards, habit, or expected sanctions rather than an orientation of trust [8, 16].

A first empirical requirement is therefore construct separation. Studies should avoid treating a single attitude item or cooperative act as an exhaustive measure of trust. A multi-method design could distinguish at least perceived trustworthiness, willingness to accept vulnerability, actual reliance, availability of safeguards, perceived exit options, and domain-specific expectations. Distrust should likewise not be inferred simply from low trust. The coexistence argument reviewed in Section 2 suggests that affirmative distrust may have its own indicators, including expectations of harmful intent, defensive monitoring, counter-verification, and interpretive suspicion [22].

The measurement of interpretive power is more difficult still. Section 6 represented effective interpretive weight as a composite of visibility, credibility, institutional position, network access, agenda-setting capacity, and consequential authority. Existing network measures can identify some forms of structural position, including centrality and brokerage [4, 7], but structural position alone is not equivalent to interpretive authority. A low-centrality expert may possess decisive evidential authority in one domain, while a highly visible actor may have limited capacity to alter formal decisions. Empirical work would therefore need to connect network position to observable changes in framing, evidence selection, downstream decisions, and the persistence of public judgments.

Re-cognizability presents a different measurement challenge because it concerns a future capability. A judgment can appear transparent at time t yet become impossible to reconstruct at $t + k$ if records are lost, categories change, reasons were never distinguished from allegations, or subsequent systems preserve only the final label. Possible indicators might include provenance completeness, accessibility of underlying evidence, traceability of revisions, existence of review triggers, and the ability of later reviewers to reproduce the procedural history of a decision. None of these is equivalent to re-cognizability by itself. They are candidate operational components that require empirical validation.

Causal identification is complicated by reciprocal influence. If trust encourages cooperation and cooperation generates evidence that increases trust, a cross-sectional association between the two does not identify direction. The same applies to status and credibility: actors may become influential because they were considered credible, while repeated deference can itself generate further credibility. A schematic system such as

$$T_t \rightarrow C_{t+1} \rightarrow H_{t+1} \rightarrow T_{t+2}$$

contains endogenous feedback by design. Empirical studies therefore require temporal information and, where feasible, interventions or natural discontinuities that help distinguish the order of effects.

Several research strategies are plausible. Longitudinal panel data can test whether prior relational states predict later changes after controlling for stable differences. Event-history designs can examine what happens around trust violations, public allegations, institutional reviews, or changes in safeguards. Network analysis can trace how interpretations and reputational signals propagate across connected actors. Experiments can isolate limited mechanisms, such as the effect of verification, reversibility, source status, or competing interpretations on willingness to cooperate, while recognizing the limited external validity of highly controlled settings. Qualitative process tracing can reconstruct how an interpretation became

authoritative, which alternatives were excluded, and how institutional records changed over time. No single method is likely to identify the full field representation.

The framework also creates a problem of scale. A relation that appears balanced at the level of formal speaking time may be highly asymmetric at the level of agenda setting; a national institution that appears trusted in aggregate may contain strong regional, class, professional, or historical differences. Aggregation can therefore produce apparent stability that does not exist locally. Empirical analysis should report the spatial, institutional, and temporal resolution at which a trust or power variable is defined and test whether conclusions survive plausible changes in that resolution.

Reflexivity produces an additional difficulty. Once indicators of trust, transparency, participation, or balance become governance targets, actors can adapt to them. A public body may optimize the visible indicator while preserving the underlying asymmetry through another channel. Metrics can therefore become part of the field they are intended to measure. This is one reason the paper avoids proposing a single public-trust index or a fixed scalar objective. Empirical operationalization should remain plural and periodically reassessed rather than treated as a neutral window onto an unchanged object.

For these reasons, the equations in the present paper should be read as hypotheses about structure rather than fitted empirical models. A stronger empirical programme would have to define measurable variables, establish discriminant validity among adjacent constructs, specify temporal and relational boundaries, and identify observable implications that could fail. Without such steps, formal elegance would risk becoming unfalsifiable description.

11.4 Scope of Normative Evaluation

The paper's normative proposals are deliberately incomplete. Sections 8–10 emphasized re-cognizability, revisability, interpretive contestability, meaningful exit, evidential responsiveness, and the preservation of alternative relational paths. These conditions address a recurrent danger identified throughout the paper: historically stabilized trust or distrust can become self-sealing and can convert provisional interpretations into durable structures of power. Yet openness to revision is not a complete theory of justice.

First, revisability can conflict with finality. Legal judgments, administrative decisions, organizational procedures, and interpersonal commitments cannot remain fully open at every moment without imposing substantial costs. Endless reconsideration can itself become a form of domination, especially when powerful actors repeatedly force vulnerable parties to relitigate settled claims. A just institution therefore needs criteria for when a matter can be reopened and when provisional closure is warranted. The framework identifies the danger of irreversible closure but does not derive the optimal threshold between finality and revision.

Second, transparency and re-cognizability can conflict with privacy, confidentiality, security, and forgetting. Preserving detailed provenance may help later reviewers reconstruct a public judgment, but it can also preserve stigmatizing information, expose protected identities, create surveillance capacity, or prevent socially valuable rehabilitation. Section 10 therefore proposed event-sensitive rather than universal recording. The appropriate balance among accountability, privacy, evidential preservation, and the possibility of forgetting remains a domain-specific normative problem. Accountability theory similarly distinguishes multiple relationships, forums, obligations, and consequences rather than treating disclosure as sufficient by itself [6].

Third, interpretive plurality is not equivalent to epistemic equality. A governance system should preserve the possibility of challenge, but it need not grant identical weight to every interpretation. Expertise, direct evidence, methodological quality, legal competence, and responsibility can justify asymmetry. Fricker's account of epistemic injustice shows why credibility distributions can be normatively distorted [10]; the converse error would be to conclude that all credibility differentials are unjust. The present paper proposes contestable asymmetry, but it does not supply a complete criterion for assigning appropriate interpretive weight.

Fourth, meaningful exit is not always available or normatively sufficient. Citizens cannot simply exit every state function; workers may face severe economic constraints; patients may have no equivalent medical alternative; members of marginalized groups may encounter the same structure across institutions.

Hirschman's distinction among exit, voice, and loyalty is useful partly because it shows that different settings make different corrective mechanisms feasible [17]. A field-governance approach should therefore avoid treating exit as a universal remedy. In some domains, the normative task is to transform the relation because exit would reproduce exclusion rather than protect agency.

Fifth, generative openness can conflict with protection against harm. A relation that remains open to every possible future trajectory may expose participants to serious and foreseeable danger. Public institutions legitimately prohibit some actions, restrict some forms of access, and stabilize some norms precisely because unrestricted generativity would be destructive. The present paper therefore does not identify justice with the maximization of possible trajectories. Its more limited proposal is that restrictions should remain identifiable, reviewable where appropriate, and responsive to relevant evidence. Substantive rights, duties, distributive principles, and prohibitions remain necessary to determine which trajectories ought to be protected or excluded.

Sixth, the evaluation of a trusting relation cannot stop at the participants inside it. Section 9 emphasized externalized harm: two actors can form a highly cooperative and mutually beneficial relation that increases their capacity to exploit, exclude, or attack others. A normative analysis must therefore specify who counts as affected by the relation and how their claims enter the evaluation. The field representation helps reveal this boundary problem but does not solve it. International relations, corporate governance, policing, platform governance, and informal organizational coalitions may each require different accounts of standing, responsibility, and permissible externalities.

Seventh, the framework remains historically and culturally underdetermined. Public expectations concerning appropriate trust, deference, privacy, expertise, collective responsibility, and legitimate authority vary across institutions and societies. A mechanism that appears paternalistic in one setting may be interpreted as care or role responsibility in another; a form of public disclosure that supports accountability in one legal culture may violate an important norm of confidentiality in another. The framework is intended to identify relational questions that travel across contexts, not to erase contextual differences in their normative answers.

These qualifications also apply to procedural legitimacy. Fair, participatory, and intelligible procedures can improve legitimacy and affect compliance [35, 11], but procedure cannot by itself redeem a substantively unjust outcome. Nor does participation guarantee that marginalized interests are effectively represented. The paper's emphasis on public re-cognizability should therefore be read as one component of legitimate governance rather than as a substitute for substantive political, legal, and ethical evaluation.

The normative contribution can consequently be stated in negative and procedural terms more confidently than in comprehensive positive terms. Trust and distrust should not become immune to relevant evidence merely because they have stabilized historically. Interpretive authority should not become impossible to contest merely because it has accumulated institutional prestige. Governance mechanisms should not be assumed legitimate merely because they were introduced to correct an earlier imbalance. And affected actors should retain, where feasible and normatively appropriate, some route by which relations can be re-examined, redirected, or institutionally reviewed.

These are boundary conditions rather than a complete theory of the good public order. They leave unresolved difficult conflicts among stability and openness, privacy and provenance, expertise and equality, protection and generativity, local cooperation and external harm, and individual exit and collective transformation. Making those conflicts explicit is preferable to resolving them implicitly through an assumption that trust itself is good. The framework developed here is therefore best treated as a preliminary diagnostic perspective: it identifies where public trust can become generative, self-reinforcing, interpretively powerful, and potentially self-sealing, while leaving substantive adjudication to further empirical and normative work.

12 Conclusion

This paper has reconsidered public trust as a historically evolving relational condition rather than as an isolated attitude, a fixed property of an institution, or a resource whose quantity should simply be maximized.

The argument has proceeded through four linked moves. It first distinguished trust, distrust, reliance, and trustworthiness while locating them within relational histories and institutional conditions. It then examined the feedback processes through which trust and distrust can become self-stabilizing, path dependent, and attractor-like. The analysis subsequently turned to interpretation, showing how stabilized trust and distrust can alter the reception of later evidence and redistribute interpretive authority. Finally, the paper introduced a preliminary field-governance representation for examining how these processes interact with status, evidence, attention, power, institutional design, and governance interventions. The resulting perspective is deliberately limited. Its contribution is not a completed theory of field governance, but a way to restate the governance of public trust as a problem of historically coupled relational conditions whose stabilization can be productive, dangerous, and revisable.

The first implication is that trust and distrust acquire significance through what they make possible in a relation. Trust can lower defensive costs, widen disclosure, support cooperation, and create room for repeated interaction. Distrust can preserve caution, motivate verification, sustain critical distance, and protect actors from premature exposure. Neither orientation, however, determines the normative quality of the trajectory that follows. Trust may support care and collective capacity, but it can also stabilize dependency, collusion, extraction, or coordinated harm toward outsiders. Distrust may protect scrutiny, but it can also reduce corrective contact, harden adversarial interpretation, and reproduce the very conditions from which further distrust emerges. The evaluative problem is therefore not exhausted by asking whether trust is high or low. It concerns the relational capacities that trust and distrust amplify, the actors who gain or lose those capacities, and the extent to which the resulting configuration remains responsive to later evidence and changing conditions.

A second implication concerns the historical reality of trust-related structures. Trust need not be treated as an independently observable substance in order to become causally consequential. Once actors organize disclosure, cooperation, interpretation, and vulnerability around a configuration understood as trust, that configuration can alter subsequent relations. Distrust can operate in the same way. An initially weak orientation can affect contact, monitoring, reputation, access to information, and the interpretation of ambiguous conduct; those effects can then produce new experiences that reinforce the original orientation. In this sense, trust and distrust can become relationally real through the trajectories they help generate. Their relational reality remains distinct from the truth of the beliefs that sustain them and from the justice of the outcomes they enable. A mistaken trust can still reorganize a relation, just as an initially weak or poorly grounded suspicion can produce a durable social structure.

This historical perspective also changes the significance of initial conditions. Public governance often encounters relations in which substantive trust is unavailable and cannot reasonably be demanded. The relevant task may then be to make limited cooperation possible under conditions of low trust by reducing the costs of error, bounding exposure, enabling verification, preserving reversibility, and creating opportunities for repeated interaction. Such arrangements are not themselves trust. They may nevertheless permit a relational history to begin. Where that history generates credible evidence of reliability, trust may later emerge as an endogenous consequence rather than as a prerequisite imposed by policy. The same reasoning applies in reverse to distrust attractors. A governance intervention need not command actors to become trusting; it can instead reopen blocked evidential channels, reduce self-reinforcing interpretive couplings, create independent review, or restore relational escape channels through which corrective experience becomes possible.

The paper has placed particular emphasis on interpretation because public trust does more than influence whether an actor is believed. Trust and distrust can change how later evidence is read, whose account is treated as presumptively credible, which ambiguities are resolved charitably or suspiciously, and which interpretations circulate widely enough to become part of public reality. Interpretive authority is therefore one of the mechanisms through which trust and power can reinforce one another. A high-status or highly connected actor may acquire greater interpretive weight; that weight can shape public responses; those responses can alter the conduct and opportunities of the actor being interpreted; and the resulting history can then appear to confirm the original account. Similar dynamics can arise around institutions, professional communities, public controversies, and organizational reputations. Publicity alone does not remove these

asymmetries. A formally open public sphere can still contain highly unequal capacities to define relevant evidence, set interpretive frames, circulate claims, and trigger institutional consequences.

The corresponding governance concern is not the elimination of interpretive asymmetry. Expertise, direct experience, institutional responsibility, and evidential access can justify unequal weight. The concern is whether such asymmetries become self-sealing: whether an interpretation gains the capacity to regulate the admission, credibility, and meaning of the very evidence by which it might otherwise be challenged. The paper has described a more defensible condition as *contestable asymmetry*. Under such a condition, unequal interpretive weight may persist, but its basis remains recoverable, its consequences remain reviewable, and alternative evidence or interpretation retains a practicable route into the public process. This shifts attention from equalizing every voice to maintaining the conditions under which authority itself can be re-evaluated.

Public re-cognizability was introduced to name one part of this requirement. A public judgment is re-cognizable when later participants can reconstruct, to an appropriate degree, how it formed: what evidence entered, how evidence and interpretation were distinguished, which actors possessed relevant authority, what alternatives were considered or excluded, and under what conditions the judgment may be reopened. Re-cognizability therefore extends beyond transparency understood as present visibility. It is historical and procedural. It concerns the preservation of paths back into the generative process of a public judgment. This does not imply universal recording, permanent contestation, or the abolition of institutional finality. Records have financial, cognitive, administrative, security, and privacy costs, and some public decisions require closure. The more limited proposal is event-sensitive: where a decision materially changes rights, access, reputation, institutional status, or future evidential position, the case for preserving sufficient provenance for later reconstruction becomes stronger.

The preliminary field-governance representation was introduced to hold these dependencies together without reducing them to a single central actor. Public trust can be influenced simultaneously by relational history, institutional safeguards, status, network position, evidence access, attention, symbolic cues, and interpretive authority. The notation Φ and the action-like functional $S[\Phi]$ provide a compact way to represent such distributed conditions and their possible coupling. A designed intervention field ϕ_e then represents a governance arrangement that perturbs the existing configuration rather than acting from outside it. The higher-order terms in the representation serve as a reminder that an intervention directed at one target can alter several others: transparency can change accountability and strategic performance; verification can change both cooperation and suspicion; public participation can expand voice while also redistributing attention and reputational power. The formalism remains heuristic until its variables, scales, and response relations are empirically specified.

This perspective also places limits on the meaning of balance. Public balance is not a state in which all actors possess equal influence, all fields vanish, or all conflict disappears. Nor is it adequately captured by one observational frame. A process can appear balanced when measured by speaking time while remaining sharply asymmetric in agenda-setting, evidence access, network reach, or the durability of reputational effects. The paper has therefore treated balance as scale-sensitive and observationally dependent. A more useful provisional criterion is whether the public configuration preserves meaningful capacities for evidential response, interpretive redirection, voice, revision, and exit across the relevant scales of analysis. Such capacities do not guarantee justice. They reduce the likelihood that a historically stabilized configuration becomes immune to the very processes through which its legitimacy and accuracy would need to be reassessed.

The normative orientation that follows is consequently one of *revisable stability*. Public relations require stabilization. Cooperation, expertise, law, organizational memory, and durable institutions would be impossible if every judgment and every trust relation were reopened from zero at every moment. Generativity therefore does not require perpetual destabilization. The problem arises when stabilization closes the channels through which the relation could later be re-cognized and revised. A generative public order can contain stable trust, durable authority, and institutional asymmetry while retaining the possibility that new evidence, new participants, externalized harms, or changed historical conditions alter the structure. In this sense, a healthy trust relation is capable of doubt, while a healthy distrust relation remains capable of

being changed by evidence.

Several research directions follow from this preliminary account. Empirical work would need to distinguish trust from reliance and distrust, identify mechanisms of relational stabilization, and examine how changes in contact, verification, reputation, and interpretive authority alter later trajectories. Network and longitudinal designs may be especially useful where public trust depends on distributed reinforcement rather than isolated judgments. Event-based analysis could examine moments at which reputational or institutional decisions substantially change later evidential possibilities. Comparative work could investigate whether similar governance arrangements produce different outcomes under different historical, cultural, or institutional fields. Formal work would need to clarify when attractor language adds explanatory value beyond path dependence or institutional persistence, and whether the field representation can generate testable response relations rather than remaining a conceptual map.

The paper's central proposal can therefore be stated narrowly. Public trust should be approached neither as an isolated psychological quantity nor as an unconditional public good. It is one component of a historically evolving relational configuration in which trust, distrust, evidence, interpretation, status, power, and institutional intervention can mutually reshape one another. Governance should accordingly attend to the conditions under which these configurations form, stabilize, and become revisable. The relevant aspiration is not maximal trust, maximal skepticism, or an imagined absence of power. It is the maintenance of public conditions in which cooperation can begin under uncertainty, authority can operate without becoming interpretively immune, distrust can remain corrigible, and stabilized judgments can still be re-entered when evidence or history gives sufficient reason. Field governance, in the limited sense developed here, is a preliminary vocabulary for studying that problem rather than a claim to have resolved it.

References

- [1] Robert Axelrod. *The Evolution of Cooperation*. New York: Basic Books, 1984. ISBN: 9780465021215.
- [2] Annette Baier. "Trust and Antitrust". In: *Ethics* 96.2 (1986), pp. 231–260. doi: 10.1086/292745.
- [3] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades". In: *Journal of Political Economy* 100.5 (1992), pp. 992–1026. doi: 10.1086/261849.
- [4] Phillip Bonacich. "Power and Centrality: A Family of Measures". In: *American Journal of Sociology* 92.5 (1987), pp. 1170–1182. doi: 10.1086/228631.
- [5] Pierre Bourdieu and Loïc J. D. Wacquant. *An Invitation to Reflexive Sociology*. Chicago: University of Chicago Press, 1992.
- [6] Mark Bovens. "Analysing and Assessing Accountability: A Conceptual Framework". In: *European Law Journal* 13.4 (2007), pp. 447–468. doi: 10.1111/j.1468-0386.2007.00378.x.
- [7] Ronald S. Burt. "Structural Holes and Good Ideas". In: *American Journal of Sociology* 110.2 (2004), pp. 349–399. doi: 10.1086/421787.
- [8] Karen S. Cook, Russell Hardin, and Margaret Levi. *Cooperation Without Trust?* New York: Russell Sage Foundation, 2005. ISBN: 9780871541642.
- [9] Claudia Eitzinger and Peter M. Wiedemann. "Trust in the Safety of Tourist Destinations: Hard to Gain, Easy to Lose? New Insights on the Asymmetry Principle". In: *Risk Analysis* 28.3 (2008), pp. 843–853. doi: 10.1111/j.1539-6924.2008.01077.x.
- [10] Miranda Fricker. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press, 2007. ISBN: 9780198237907. doi: 10.1093/acprof:oso/9780198237907.001.0001.
- [11] Archon Fung. "Varieties of Participation in Complex Governance". In: *Public Administration Review* 66.s1 (2006), pp. 66–75. doi: 10.1111/j.1540-6210.2006.00667.x.

- [12] Martin Gargiulo and Mario Benassi. "Trapped in Your Own Net? Network Cohesion, Structural Holes, and the Adaptation of Social Capital". In: *Organization Science* 11.2 (2000), pp. 183–196. doi: 10.1287/orsc.11.2.183.12514.
- [13] Nicole Gillespie and Graham Dietz. "Trust Repair after an Organization-Level Failure". In: *Academy of Management Review* 34.1 (2009), pp. 127–145. doi: 10.5465/amr.2009.35713319.
- [14] Mark Granovetter. "Threshold Models of Collective Behavior". In: *American Journal of Sociology* 83.6 (1978), pp. 1420–1443. doi: 10.1086/226707.
- [15] Mark S. Granovetter. "The Strength of Weak Ties". In: *American Journal of Sociology* 78.6 (1973), pp. 1360–1380. doi: 10.1086/225469.
- [16] Russell Hardin. *Trust and Trustworthiness*. New York: Russell Sage Foundation, 2002. ISBN: 9780871543417.
- [17] Albert O. Hirschman. *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*. Cambridge, MA: Harvard University Press, 1970. ISBN: 9780674276604.
- [18] Robert Jervis. "Cooperation under the Security Dilemma". In: *World Politics* 30.2 (1978), pp. 167–214. doi: 10.2307/2009958.
- [19] Karen Jones. "Trust as an Affective Attitude". In: *Ethics* 107.1 (1996), pp. 4–25. doi: 10.1086/233694.
- [20] Peter H. Kim et al. "Removing the Shadow of Suspicion: The Effects of Apology Versus Denial for Repairing Competence- Versus Integrity-Based Trust Violations". In: *Journal of Applied Psychology* 89.1 (2004), pp. 104–118. doi: 10.1037/0021-9010.89.1.104.
- [21] Margaret Levi and Laura Stoker. "Political Trust and Trustworthiness". In: *Annual Review of Political Science* 3 (2000), pp. 475–507. doi: 10.1146/annurev.polisci.3.1.475.
- [22] Roy J. Lewicki, Daniel J. McAllister, and Robert J. Bies. "Trust and Distrust: New Relationships and Realities". In: *Academy of Management Review* 23.3 (1998), pp. 438–458. doi: 10.5465/amr.1998.926620.
- [23] Kurt Lewin. *Field Theory in Social Science: Selected Theoretical Papers*. Ed. by Dorwin Cartwright. New York: Harper, 1951.
- [24] Charles G. Lord, Lee Ross, and Mark R. Lepper. "Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence". In: *Journal of Personality and Social Psychology* 37.11 (1979), pp. 2098–2109. doi: 10.1037/0022-3514.37.11.2098.
- [25] Niklas Luhmann. *Trust and Power: Two Works*. Trans. by Howard Davis, John Raffan, and Kathryn Rooney. Chichester: John Wiley & Sons, 1979. ISBN: 0471997587.
- [26] Roger C. Mayer, James H. Davis, and F. David Schoorman. "An Integrative Model of Organizational Trust". In: *Academy of Management Review* 20.3 (1995), pp. 709–734. doi: 10.5465/AMR.1995.9508080335.
- [27] Elinor Ostrom. "A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997". In: *American Political Science Review* 92.1 (1998), pp. 1–22. doi: 10.2307/2585925.
- [28] Elinor Ostrom. *Understanding Institutional Diversity*. Princeton, NJ: Princeton University Press, 2005. doi: 10.2307/j.ctt7s7wm.
- [29] Kenneth A. Oye. "Explaining Cooperation under Anarchy: Hypotheses and Strategies". In: *World Politics* 38.1 (1985), pp. 1–24. doi: 10.2307/2010349.
- [30] Paul Pierson. "Increasing Returns, Path Dependence, and the Study of Politics". In: *American Political Science Review* 94.2 (2000), pp. 251–267. doi: 10.2307/2586011.
- [31] Roberta Sellaro et al. "A Question of Scent: Lavender Aroma Promotes Interpersonal Trust". In: *Frontiers in Psychology* 5 (2015), p. 1486. doi: 10.3389/fpsyg.2014.01486.

- [32] Denise Skinner, Graham Dietz, and Antoinette Weibel. “The Dark Side of Trust: When Trust Becomes a ‘Poisoned Chalice’”. In: *Organization* 21.2 (2014), pp. 206–224. doi: 10.1177/1350508412473866.
- [33] Paul Slovic. “Perceived Risk, Trust, and Democracy”. In: *Risk Analysis* 13.6 (1993), pp. 675–682. doi: 10.1111/j.1539-6924.1993.tb01329.x.
- [34] Charles S. Taber and Milton Lodge. “Motivated Skepticism in the Evaluation of Political Beliefs”. In: *American Journal of Political Science* 50.3 (2006), pp. 755–769. doi: 10.1111/j.1540-5907.2006.00214.x.
- [35] Tom R. Tyler. *Why People Obey the Law*. New Haven: Yale University Press, 1990. ISBN: 9780300044034.
- [36] Brian Uzzi. “Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness”. In: *Administrative Science Quarterly* 42.1 (1997), pp. 35–67. doi: 10.2307/2393808.