The Restoration Matrix: Six Falsifiable Operators for Governance Coherence

Allison Hensgen* Joel Thorarinson†

June 2026

Abstract

We propose six governance coherence operators, each with a pre-registered falsification protocol. Governance systems optimized for procedural compliance routinely produce social fragmentation, ecological degradation, and institutional distrust — a gap that existing governance theory documents but does not operationalize. Building on the coherence operator $\Delta = (P \cdot A \cdot R)/(D + N)$ introduced in our prior work on system drift detection [Thorarinson and Hensgen, 2026], we formalize six *restoration operators* Δ_1 through Δ_6 that decompose governance coherence into independently measurable components: collective physiological synchrony (Δ_1), narrative institutional memory (Δ_2), transparency and social trust (Δ_3), ecological grounding (Δ_4), institutional restraint (Δ_5), and deliberative environment quality (Δ_6). Each operator is linked to specific measurement protocols using validated instruments — heart rate variability coherence and inter-brain EEG synchrony [Palumbo et al., 2017, Vickhoff et al., 2013], institutional trust scales [Mayer et al., 1995, Edelman Trust Institute, 2025], deliberative quality indices [Fishkin, 2009], and social cohesion measures [Chan et al., 2006]. The primary contribution is the falsification architecture: for each operator, we specify an experimental design with *a priori* power analysis, predicted effect sizes grounded in existing meta-analytic estimates, and an explicit criterion under which the operator would be rejected. The composite governance coherence score $G_c(t)$ integrates the six operators into a single diagnostic metric. The framework does not replace formal governance structures but adds a diagnostic layer designed to detect coherence degradation before institutional failure becomes visible in conventional metrics. If governance systems scoring high on $G_c(t)$ do not outperform those scoring low on measures of trust, participation, and institutional resilience, the framework is wrong.

Keywords: governance coherence; falsification protocols; institutional trust; deliberative quality; social capital; commons governance; restoration operators; coherence measurement

1 Introduction

Modern governance systems are optimized for procedural correctness, legal compliance, auditability, and hierarchical accountability. These properties are necessary. They are not sufficient. A governance system can be procedurally correct and socially distrusted. It can be legally compliant and ecologically extractive. It can be auditable and alienating to the populations it governs. The gap between procedural adequacy and functional governance is a structural problem with measurable consequences — and, crucially, one that most governance measurement frameworks leave unmeasured.

The evidence for this gap is extensive. Democracies with intact legal frameworks experience declining civic participation and rising institutional distrust [Putnam, 2000]; the 2025 Edelman Trust Barometer reports that five of the ten largest global economies now rank among the least-trusting nations, with trust indices as low as 37% [Edelman Trust Institute, 2025]. Commons

*Coherence Research Group. ORCID: 0009-0008-7247-0307

†Coherence Research Group. ORCID: 0000-0002-0553-842X. joel.thorarinson@conformalmaps.com

governance succeeds or fails based on conditions — face-to-face communication, shared identity, graduated sanctions, conflict resolution mechanisms — that are not reducible to the formal rules of the system [Ostrom, 1990]. Organizations with identical formal structures exhibit dramatically different levels of trust, resilience, and adaptive capacity [Mayer et al., 1995]. The Worldwide Governance Indicators [Kaufmann et al., 2010] capture six dimensions of institutional quality, and the Varieties of Democracy project [Coppedge et al., 2022] measures deliberative quality and participatory depth, yet neither operationalizes the structural conditions that produce or erode institutional coherence over time. Complex adaptive systems in governance exhibit nonlinear dynamics, threshold effects, and limited predictability that conventional governance theory acknowledges but struggles to operationalize [Duit and Galaz, 2008].

We propose that the missing explanatory variable is *coherence infrastructure*: the set of conditions that determine whether a governance system's components — its people, processes, institutions, narratives, and physical environments — maintain coordinated, recoverable behavior over time. In our prior work [Thorarinson and Hensgen, 2026], we formalized coherence as a measurable property of complex systems using operators drawn from dynamical systems theory and recurrence quantification analysis. We demonstrated that coherence degradation precedes conventional failure indicators by a factor of 2.2× on average across seven domains, from turbofan engines to cardiac rhythms.

The present paper extends this framework to governance. We identify six structural conditions — which we call *restoration operators* — that constitute the coherence infrastructure of a governance system. Each operator is defined with a formal functional form, connected to existing empirical research, and linked to specific measurement protocols using validated instruments. However, the primary contribution is not the operators themselves but their **falsification architecture**. Most governance measurement frameworks — including social capital indices, deliberative quality scores, and institutional trust surveys — propose measures without specifying what would count as evidence against them. For each of our six operators, we provide an explicit experimental design with *a priori* power analysis, predicted effect sizes grounded in existing meta-analytic estimates, control conditions, and a falsification criterion: a specific outcome that, if observed, would require rejecting the operator. If a governance system scores high on all six operators and still fails, the framework is wrong.

1.1 Contributions

1. We formalize **governance coherence** as an operational construct, distinct from compliance, legitimacy, or efficiency, and show how the coherence operator Δ applies to institutional systems.

2. We define **six restoration operators** Δ_1 through Δ_6 — collective physiological synchrony, narrative institutional memory, transparency and social trust, ecological grounding, institutional restraint, and deliberative environment quality — each with a formal functional form, measurable inputs, and empirical grounding.

3. We derive a **composite governance coherence score** $G_c(t)$ that integrates the six operators into a single diagnostic metric.

4. We present **six falsification protocols** (constituting the core contribution) with explicit experimental designs, sample sizes derived from *a priori* power analyses, control conditions, predicted effect sizes from existing meta-analytic data, and falsification criteria specifying conditions under which each operator would be rejected.

2 Theoretical Foundation

2.1 Coherence as a Governance Condition

In Thorarinson and Hensgen [2026], we defined the coherence score for a complex system as:

$$\Delta(t) = \frac{P(t) \cdot A(t) \cdot R(t)}{D(t) + N(t) + \epsilon} \tag{1}$$

where P measures pattern retention (does the system revisit stable states?), A measures phase alignment (do the components synchronize?), R measures recovery capacity (can the system return to its attractor after perturbation?), D measures drift (systematic departure from baseline), and N measures noise amplification (does the system amplify or dampen stochastic perturbations?).

In the governance context, these operators acquire institutional meaning:

- **Pattern retention** P: Does the institution maintain its core behavioral patterns across leadership transitions, electoral cycles, and external shocks? An institution that reinvents itself every cycle has low P.
- **Phase alignment** A: Is there alignment between the institution's stated values and its observable behavior? Between its different branches, levels, or stakeholder groups? Misalignment between rhetoric and action is measurable desynchronization.
- **Recovery capacity** R: Can the institution recover from crises — scandals, financial shocks, leadership failures — and return to functional operation? Or does each crisis leave permanent damage?
- **Drift** D: Is the institution's actual behavior drifting from its founding mission, its legal mandate, or its stakeholders' expectations? Mission drift is the governance analog of signal drift in engineered systems.
- **Noise amplification** N: Does the institution amplify social noise — rumors, polarization, factional conflict — or dampen it? A governance system that amplifies noise is structurally unstable regardless of its formal procedures.

This mapping shares the same mathematical structure, though the governance-specific operators require independent empirical validation. The governance application connects to Holling [2001]'s panarchy model, in which hierarchical adaptive cycles at different scales determine system resilience through the interplay of exploitation, conservation, release, and reorganization phases. Governance coherence degrades when stabilizing forces (P, A, R) weaken relative to destabilizing forces (D, N). The question is what *produces* the stabilizing forces in a governance context. That is what the six restoration operators address.

The construct validity of the composite measure $G_c(t)$ follows the framework of Cronbach and Meehl [1955]: each operator must demonstrate convergent validity with established measures of institutional function and discriminant validity from adjacent constructs like procedural compliance or political approval ratings.

2.2 The Measurement Problem: Assigning $[0, 1]$ Scores to Qualitative Constructs

A central challenge for any governance coherence framework is the mapping from qualitative concepts — trust, narrative coherence, deliberative quality — to interval-scale scores on $[0, 1]$. We address this challenge through three strategies.

Strategy 1: Validated survey instruments. Where validated psychometric instruments exist, we use them directly. Institutional trust is measured via the Organizational Trust Inventory [Mayer et al., 1995] or equivalent (e.g., the Edelman Trust Barometer's methodology for population-level trust; Edelman Trust Institute 2025). Social cohesion is measured via the six-dimension framework of Chan et al. [2006], which provides validated survey items for social relations, task relations, perceived unity, emotions, social inclusion, and equality. Deliberative quality is assessed via Fishkin's Deliberative Polling protocol [Fishkin, 2009], which yields interval-scale measures of opinion quality, information gain, and preference structure. Occupant affect is measured via the Positive and Negative Affect Schedule (PANAS; Watson et al. 1988). In each case, raw scores are min-max normalized to $[0, 1]$ within the instrument's established range.

Strategy 2: Physiological measurements. Where physiological measurements are available, they provide continuous-scale data that map naturally to $[0, 1]$. Heart rate variability (HRV) coherence [Palumbo et al., 2017] is the exemplar: Vickhoff et al. [2013] demonstrated that choral singing entrains HRV across participants, and the HRV coherence ratio (proportion of spectral power in the 0.04–0.15 Hz band) yields a continuous score that can be normalized against population baselines. Inter-brain EEG synchrony [Dumas et al., 2010, Reinero et al., 2021] provides a second physiological channel, with phase-locking values yielding $[0, 1]$ scores directly. We designate HRV coherence as the primary bio-measurement exemplar for this framework because of its non-invasiveness (wearable monitors), established measurement validity [Palumbo et al., 2017], and demonstrated sensitivity to social coordination [Vickhoff et al., 2013, McCraty et al., 2009].

Strategy 3: Structured behavioral coding. For constructs where neither validated surveys nor physiological measures are sufficient — notably boundary stability (B_s) in Δ_5 and environmental stability (E_s) in Δ_6 — we propose structured behavioral coding by blinded observers using pre-registered codebooks. The proportion of coded events meeting criteria yields a $[0, 1]$ score directly. Inter-rater reliability (Cohen's $\kappa > 0.7$) is required before scores are used.

We acknowledge that the choice of normalization method (min-max vs. z-score vs. percentile rank) affects the composite score $G_c(t)$, and that cross-cultural calibration of survey instruments introduces measurement invariance concerns. These issues are addressed in the Limitations section. The falsification protocols in Section 5 are designed to be robust to normalization choice: each protocol tests a between-condition difference, so monotonic transformations of the $[0, 1]$ scale do not affect the falsification criterion.

3 The Six Restoration Operators

Each operator is defined as a function of measurable inputs. The functional forms are not arbitrary — they reflect the multiplicative interaction of components that are jointly necessary (if any input approaches zero, the operator degrades) while remaining independently measurable. For each operator, we specify the measurement instruments and scoring procedures that map observable quantities to the $[0, 1]$ inputs defined in the formal expressions.

3.1 Δ_1: Collective Physiological Synchrony

Definition 1 (Synchrony Operator)**.** *The collective physiological synchrony operator measures the degree of physiological and behavioral synchrony among governance participants:*

$$\Delta_1 = f(S_b, R_h, M_c, G_p) = \frac{S_b \cdot R_h \cdot M_c \cdot G_p}{\sigma_{\Delta_1} + \epsilon} \quad (2)$$

where $S_b \in [0, 1]$ is breath and physiological synchrony, $R_h \in [0, 1]$ is rhythmic entrainment, $M_c \in [0, 1]$ is movement coordination, $G_p \in [0, 1]$ is group presence (attentional co-orientation), and σ_{Δ_1} is the variance of the operator over the measurement window.

What it measures. Groups that move together, breathe together, and attend to the same focal points develop measurably higher cooperation and trust. This is not speculative. Wiltermuth and Heath [2009] demonstrated that synchronous movement increases subsequent cooperative behavior in economic games, with effect sizes ranging from $d = 0.48$ to $d = 0.74$ across three experiments. Two independent meta-analyses confirm the robustness of this finding: Vicaria and Dickens [2017] synthesized 88 effect sizes across 42 studies ($N > 4{,}000$) and found a reliable positive effect of synchrony on prosocial behavior, while Mogan et al. [2017] reported a pooled effect of $d = 0.58$ (95% CI: 0.40–0.77) for cooperative behavior following synchronous activity. Dumas et al. [2010] showed that inter-brain EEG synchronization during social interaction is both measurable and predictive of interaction quality. Reinero et al. [2021] extended this to team settings, demonstrating that inter-brain synchrony in groups of four predicted collective performance ($N = 174$, 44 groups). McCraty et al. [2009] documented that heart rate variability (HRV) coherence between individuals correlates with reported empathy and social bonding, and Vickhoff et al. [2013] showed that choral singing entrains HRV across participants, linking shared rhythmic activity to physiological synchronization. Palumbo et al. [2017] provide a comprehensive systematic review of interpersonal autonomic physiology, establishing the measurement validity of physiological synchrony as a construct.

The institutional and sociological literature provides additional grounding. Durkheim [1912] identified "collective effervescence" — the heightened emotional and cognitive state produced by shared ritual — as the mechanism through which social solidarity is generated and maintained, a construct that maps directly to Putnam [2000]'s social capital. Rappaport [1999] argued that ritual is constitutive of institutional legitimacy, not merely an expression of it. Sosis and Alcorta [2003] formalized this in terms of costly signaling theory, showing that ritual participation serves as a credible signal of group commitment, with communities featuring more demanding rituals exhibiting higher levels of intragroup cooperation — a finding with direct implications for the design of governance deliberation processes.

Governance application. Governance spaces that include shared ritual — opening ceremonies, moments of silence, synchronized deliberation protocols — produce measurably different outcomes than those that proceed directly to agenda items. The operator does not prescribe specific rituals; it measures whether the governance process produces physiological synchrony among participants, by whatever means.

Measurement protocol. We designate HRV coherence as the primary physiological measure for Δ_1 because of three properties: (1) non-invasiveness — wearable chest-strap or wrist-based monitors can be deployed in field governance settings without disrupting proceedings; (2) established measurement validity — Palumbo et al. [2017] systematically reviewed the interpersonal autonomic physiology literature and confirmed that inter-personal HRV synchrony is a reliable, replicable construct; and (3) demonstrated sensitivity to social coordination — Vickhoff et al. [2013] showed that HRV entrainment during choral singing tracks the rhythmic structure of the shared activity with temporal precision. The HRV coherence ratio (proportion of spectral power in the 0.04–0.15 Hz band) yields a continuous $[0, 1]$ score normalizable against population baselines.

Secondary measures include inter-brain EEG synchrony via portable hyperscanning systems (higher measurement precision but more intrusive), movement coordination via accelerometer data or video-based pose estimation, and group attentional convergence via eye-tracking.

3.2 Δ_2: Narrative Institutional Memory

Definition 2 (Narrative Coherence Operator)**.** *The narrative institutional memory operator measures the strength and transmission fidelity of the governance system's founding narrative:*

$$\Delta_2 = f(N_s, I_c, M_r, C_t) = \frac{N_s \cdot I_c \cdot M_r \cdot C_t}{\sigma_{\Delta_2} + \epsilon} \tag{3}$$

where $N_s \in [0,1]$ is narrative strength (clarity and emotional resonance of the institutional story), $I_c \in [0,1]$ is identity coherence (degree to which participants identify with the narrative), $M_r \in [0,1]$ is memory retention (how well participants can articulate the narrative), and $C_t \in [0,1]$ is cultural transmission (rate at which new members acquire the narrative).

What it measures. Every durable institution maintains a coherent account of its origins, purpose, and trajectory — what organizational scholars call institutional narrative and what political scientists recognize as a component of legitimacy [Beetham, 1991]. When this narrative is clear, shared, and broadly endorsed, participants orient their behavior toward shared goals without requiring constant top-down direction. When the narrative fragments — when different factions tell incompatible stories, or when no one can articulate the institutional purpose at all — coordination costs increase and trust degrades. Bruner [1991] established that narrative is the primary cognitive structure through which humans organize experience and construct meaning. The legitimacy of governance institutions depends not only on procedural correctness but on the perceived coherence of the institutional narrative.

Governance application. A city government whose employees can articulate why the city exists and what it is trying to become will outperform one whose employees describe their work in purely procedural terms. A commons governance system whose members share a narrative about their relationship to the resource will manage it more sustainably than one whose members view participation as transactional [Ostrom, 1990]. Chan et al. [2006]'s "perceived unity" and "social relations" dimensions of social cohesion provide validated cross-sectional measures of exactly this construct.

Measurement protocol. Narrative strength (N_s) via computational text analysis of institutional communications (lexical diversity, coherence scores, sentiment consistency). Identity coherence (I_c) via the social identification subscale from Chan et al. [2006]'s social cohesion framework or equivalent organizational identification instruments. Memory retention (M_r) via structured interviews testing participants' ability to articulate the institutional narrative, scored by blinded coders against a pre-registered codebook. Cultural transmission (C_t) via longitudinal tracking of narrative acquisition among new members at 1, 3, and 6 months post-entry.

3.3 Δ_3: Public Visibility and Social Trust

Definition 3 (Transparency Operator)**.** *The public visibility operator measures the degree to which governance processes are visible, auditable, and perceived as legitimate:*

$$\Delta_3 = f(T_v, A_o, W_p, L_s) = \frac{T_v \cdot A_o \cdot W_p \cdot L_s}{\sigma_{\Delta_3} + \epsilon} \tag{4}$$

where $T_v \in [0,1]$ is visible transparency (are decisions and their rationales publicly accessible?), $A_o \in [0,1]$ is audit openness (can stakeholders verify claims independently?), $W_p \in [0,1]$ is public witnessing (do stakeholders observe the governance process directly?), and $L_s \in [0,1]$ is legitimacy signal (do stakeholders report the process as legitimate?).

What it measures. Transparency is not merely a normative ideal; it is a measurable condition with causal effects on trust and participation. Grimmelikhuijsen et al. [2013] demonstrated in a cross-national experiment ($N = 1{,}010$ across the Netherlands and South Korea) that government transparency has significant effects on perceived trustworthiness, though the relationship is moderated by cultural context and prior trust levels — a finding consistent with cross-national governance indicator data from the World Governance Indicators [Kaufmann et al., 2010] and the Varieties of Democracy project [Coppedge et al., 2022]. The operator captures not just whether information is available but whether stakeholders can verify it independently (A_o) and whether they actually observe the process (W_p). Fung [2006] provides a taxonomy of participatory mechanisms — the "democracy cube" — that maps communication intensity, participant selection, and authority levels, offering validated categories for measuring W_p and L_s.

Governance application. Open-data dashboards, public deliberation, live-streamed hearings, and independent audit access all contribute to Δ_3. But visibility without verifiability (T_v high, A_o low) produces performative transparency that can reduce trust rather than increase it. The multiplicative structure of the operator captures this: all four components must be present for the operator to score high.

Measurement protocol. Visible transparency via information accessibility audits. Audit openness via independent verification exercises. Public witnessing via attendance and engagement metrics for public proceedings. Legitimacy signal via validated trust survey instruments [Mayer et al., 1995].

3.4 Δ_4: Ecological Grounding

Definition 4 (Biophysical Accountability Operator)**.** *The ecological grounding operator measures the degree to which governance decisions are constrained by and accountable to biophysical reality:*

$$\Delta_4 = f(E_r, B_m, R_g, L_e) = \frac{E_r \cdot B_m \cdot R_g \cdot L_e}{\sigma_{\Delta_4} + \epsilon} \tag{5}$$

where $E_r \in [0,1]$ *is ecological restoration (is the governed territory's ecological condition improving?),* $B_m \in [0,1]$ *is biophysical metrics integration (are ecological measurements incorporated into governance decisions?),* $R_g \in [0,1]$ *is regenerative grounding (does the governance system aim for regeneration rather than mere sustainability?), and* $L_e \in [0,1]$ *is local ecological accountability (are governance actors accountable for ecological outcomes within their jurisdiction?).*

What it measures. Governance systems that ignore their ecological substrate eventually face crises — water shortages, soil depletion, biodiversity collapse — that destabilize the social systems they govern. This is not an environmental argument appended to governance theory; it is a structural observation about the conditions under which governance systems maintain coherence over time [Folke et al., 2010, Folke, 2006]. The resilience literature demonstrates that social-ecological systems exhibit adaptive cycles at multiple scales [Walker et al., 2004, Holling, 2001], and that governance systems embedded within degrading ecological substrates lose the adaptive capacity required for long-term institutional survival. Biermann et al. [2012] argue that the Anthropocene demands a fundamental restructuring of Earth system governance to integrate biophysical accountability into institutional decision-making. A governance system that depletes its ecological base is in the same structural position as an engine that is losing bearing lubrication: the failure is predictable even if the timing is not.

Governance application. Ecological grounding means that budgets, land-use decisions, infrastructure investments, and regulatory frameworks are informed by and accountable to biophysical measurements. This includes soil carbon as a metric for agricultural governance, water quality indices for watershed management, biodiversity surveys for land-use planning, and atmospheric monitoring for energy policy.

Measurement protocol. Ecological restoration via longitudinal ecological surveys (soil carbon, water quality, species counts). Biophysical metrics integration via content analysis of governance decisions for ecological data citations. Regenerative grounding via policy analysis comparing sustainability vs. regeneration framing. Local ecological accountability via tracking whether governance actors face consequences for ecological degradation [Walker and Salt, 2012].

3.5 Δ_5: Institutional Restraint and Deliberative Quality

Definition 5 (Power Constraint Operator)**.** *The institutional restraint operator measures the degree to which governance structures constrain the exercise of power, enforce deliberative norms, and de-escalate conflict:*

$$\Delta_5 = f(H_p, E_r, C_d, B_s) = \frac{H_p \cdot E_r \cdot C_d \cdot B_s}{\sigma_{\Delta_5} + \epsilon} \tag{6}$$

where $H_p \in [0,1]$ is deliberative pause (are there structural mechanisms requiring decision-makers to pause, consult, and consider alternatives before acting?), $E_r \in [0,1]$ is ethical restraint (do actors refrain from exercising power they legally could?), $C_d \in [0,1]$ is conflict de-escalation (does the system reduce rather than amplify interpersonal and factional conflict?), and $B_s \in [0,1]$ is boundary stability (are the boundaries of authority respected and maintained?).

What it measures. Governance systems fail when actors maximize the power available to them rather than exercising restraint. Owens et al. [2013] demonstrated across eight lab and field samples that expressed humility in organizational leaders — defined as acknowledging limitations, spotlighting others' strengths, and modeling teachability — predicts team performance, engagement, and learning orientation, with expressed humility explaining incremental variance in performance beyond conscientiousness, self-efficacy, and general mental ability. The operator measures whether the governance system structurally constrains power exercise, not whether individual actors happen to be humble. Deliberative democracy theory provides normative grounding: Habermas [1996] argues that legitimate law must be traceable to discursive processes in which all affected parties can participate, and Fishkin [2009] demonstrates through deliberative polling experiments that structured deliberation with information provision, small-group discussion, and plenary sessions produces measurable opinion change ($d = 0.30$–0.50) and higher-quality preference formation compared to unstructured polling.

Governance application. Cooling-off periods before major decisions, mandatory consultation requirements, sunset clauses, ombudsman offices, and graduated response protocols all contribute to Δ_5. A governance system that allows rapid, unilateral, irreversible decisions scores low on deliberative pause (H_p) regardless of the character of its current leaders.

Measurement protocol. Deliberative pause (H_p) via structural analysis of decision-making procedures (mandatory pause requirements, consultation mandates, cooling-off periods — scored as proportion of decision points with enforced pauses). Ethical restraint (E_r) via tracking decisions where actors chose not to exercise available power, scored as proportion of available-power-exercise opportunities that were declined. Conflict de-escalation (C_d) via conflict event frequency and severity metrics, coded by blinded observers. Boundary stability (B_s) via boundary violation tracking. Physiological validation via HRV coherence measurements during governance decision-making sessions.

3.6 Δ_6: Deliberative Environment Quality

Definition 6 (Environmental Coherence Operator)**.** *The deliberative environment quality operator measures the degree to which the physical and sensory environment of governance supports coherent cognitive and deliberative functioning:*

$$\Delta_6 = f(D_q, S_h, E_s, A_f) = \frac{D_q \cdot S_h \cdot E_s \cdot A_f}{\sigma_{\Delta_6} + \epsilon} \tag{7}$$

where $D_q \in [0,1]$ is design quality (do governance spaces meet established environmental design standards, scored via validated post-occupancy evaluation instruments?), $S_h \in [0,1]$ is sensory suitability (are lighting, acoustics, temperature, and air quality within ranges known to support sustained cognitive performance?), $E_s \in [0,1]$ is environmental stability (is the physical environment maintained and consistent across sessions?), and $A_f \in [0,1]$ is occupant affect (do participants report positive affect in the space, measured via the Positive and Negative Affect Schedule, PANAS; Watson et al. 1988).

What it measures. The physical environment in which governance occurs is not neutral. Kellert [2008] documented that biophilic design — incorporating natural materials, daylight, vegetation, and water features — measurably improves occupant well-being, cognitive function, and productivity. Controlled studies of workplace environments report that access to natural daylight produces 10–15% increases in cognitive task performance, with measurable improvements in sustained attention, working memory, and creative problem-solving. The operator does not invoke subjective aesthetic judgment; it measures environmental conditions against validated standards and captures occupant affect via standardized instruments.

Operationalization challenge. Design quality (D_q) is the most difficult component to operationalize. We propose scoring it via established post-occupancy evaluation (POE) instruments that assess spatial quality, thermal comfort, visual environment, and acoustic conditions against normative benchmarks. The WELL Building Standard and LEED Indoor Environmental Quality credits provide validated rubrics that yield interval-scale scores. Where no validated instrument exists for a specific governance space type, we recommend piloting the Building Use Studies (BUS) methodology, which produces [0,1]-normalized satisfaction scores across 12 environmental dimensions and has been validated across >800 buildings internationally.

Governance application. Council chambers, courtrooms, public hearing rooms, and administrative offices are governance infrastructure as much as the legal codes enforced within them. The quality of the physical environment affects attention, affect, and decision quality. This is not an argument for luxury; it is an argument for environments designed to support human cognitive function — natural light within 300–500 lux for reading tasks, background noise below 40 dB(A), CO_2 below 1000 ppm, and temperatures between 20–23°C.

Measurement protocol. Design quality via post-occupancy evaluation instruments (BUS, WELL, or equivalent). Sensory suitability via continuous environmental sensor data (lux meters, sound level meters, temperature and CO_2 sensors). Environmental stability via maintenance logs and condition assessments comparing cross-session variance. Occupant affect via the PANAS [Watson et al., 1988] administered pre- and post-session.

4 Integrated Model

The six restoration operators combine into a composite governance coherence score:

Definition 7 (Governance Coherence Score)**.**

$$G_c(t) = \frac{R_s(t) \cdot N_c(t) \cdot T_p(t) \cdot E_g(t) \cdot H_e(t) \cdot B_f(t)}{D_g(t) + N_g(t) + \epsilon} \tag{8}$$

where $R_s = \Delta_1$ *(collective physiological synchrony),* $N_c = \Delta_2$ *(narrative institutional memory),* $T_p = \Delta_3$ *(transparency and social trust),* $E_g = \Delta_4$ *(ecological grounding),* $H_e = \Delta_5$ *(institutional restraint),* $B_f = \Delta_6$ *(deliberative environment quality),* D_g *is governance drift (systematic departure from institutional mission), and* N_g *is social noise (polarization, fragmentation, rumor amplification).*

Figure 1 illustrates three hypothetical governance system profiles — healthy, degrading, and failing — plotted across the six operators. The multiplicative structure of the numerator is deliberate: it encodes the hypothesis that the six operators are jointly necessary. A governance system with strong synchrony, narrative coherence, transparency, ecological grounding, and institutional restraint but conducting deliberations in an environment that undermines cognitive function ($B_f \to 0$) will experience coherence degradation. A system with all six operators strong but experiencing rapid mission drift (D_g large) will score low despite its structural assets.

Proposition 1 (Coherence Degradation Precedes Institutional Failure)**.** *Under a model of gradual institutional degradation where one or more restoration operators decline at rate* γ*, the governance coherence score* $G_c(t)$ *crosses the critical threshold* $G_c^* = 0.3$ *at time* t_c *satisfying:*

$$t_c \leq t_f - \frac{1}{\gamma} \log\left(\frac{F_{threshold}}{\sigma_{G_c}}\right) \tag{9}$$

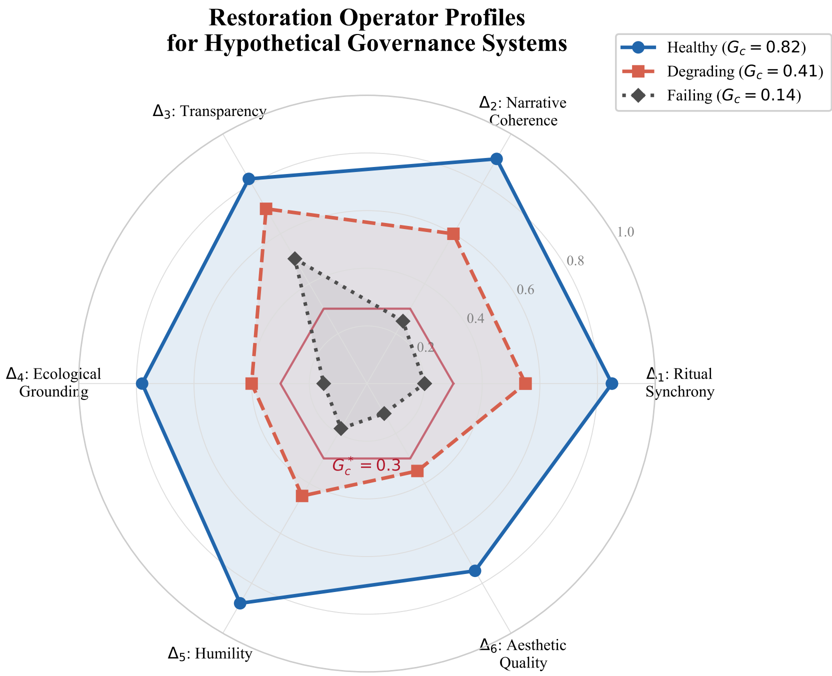


Figure 1: Restoration operator profiles for three hypothetical governance systems. The *healthy* system (solid blue, $G_c = 0.82$) scores above 0.7 on all six operators. The *degrading* system (dashed red, $G_c = 0.41$) shows selective decline in ecological grounding (Δ_4), institutional restraint (Δ_5), and deliberative environment quality (Δ_6). The *failing* system (dotted grey, $G_c = 0.14$) falls below the critical threshold $G_c^* = 0.3$ (red circle) on all operators. The multiplicative structure of G_c means that collapse of any single operator degrades the composite score.

where t_f is the time of observable institutional failure and $F_{threshold}$ is the sensitivity of conventional failure metrics (voter turnout decline, trust survey drops, staffing crises). Since coherence operators measure structural conditions rather than outcome metrics, $\sigma_{G_c} \ll F_{threshold}$ in general, yielding $t_c \ll t_f$.

This proposition is the governance analog of the early detection result in Thorarinson and Hensgen [2026]: coherence degradation is detectable before outcome-level failure metrics trigger, for the same mathematical reason — structural measures are finer invariants than aggregate statistics. Figure 2 illustrates this lead-time advantage over a hypothetical 48-month political cycle.

5 Falsification Protocols

This section constitutes the primary contribution of the paper. A framework that cannot be falsified is not a scientific framework [Popper, 1959]. Most governance measurement proposals — including prominent indices of institutional quality, social capital, and deliberative democracy — lack explicit falsification criteria. They propose measures without specifying what evidence would count against them. We take the opposite approach: for each of the six restoration operators, we provide an explicit experimental design with (1) a directional hypothesis derived from the operator's theoretical claims, (2) a control condition, (3) an *a priori* sample size justified by existing meta-analytic effect size estimates, (4) predicted outcomes with specific effect sizes, and (5) a falsification criterion — a specific result that, if observed, requires rejecting the operator

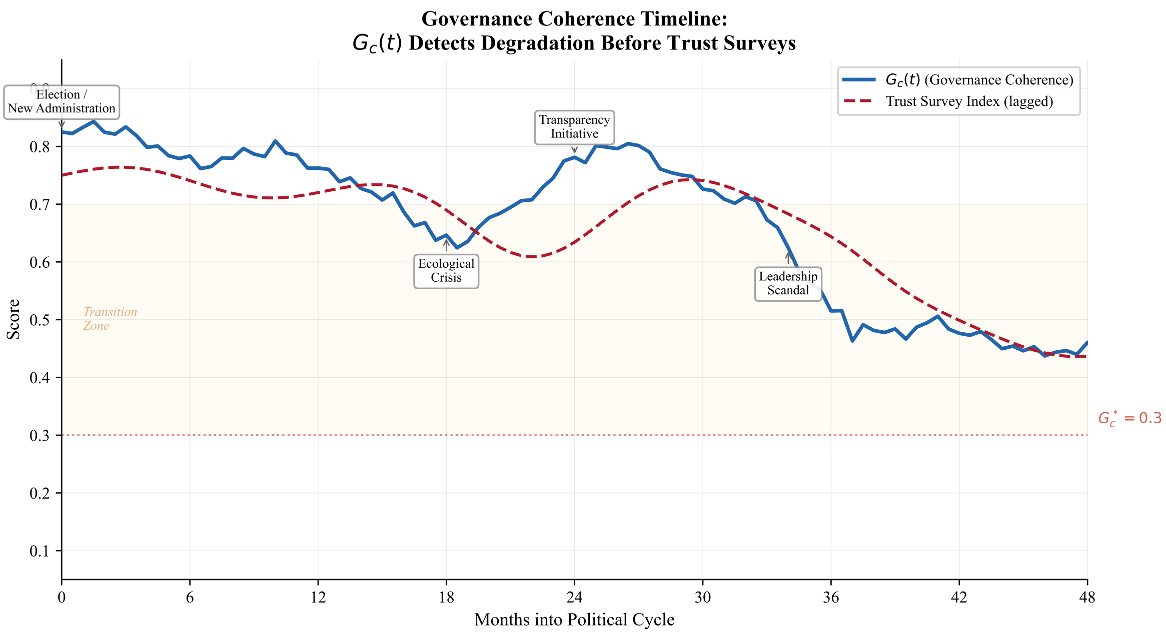


Figure 2: Governance coherence $G_c(t)$ (solid blue) and a lagged trust survey index (dashed red) over a 48-month political cycle. G_c detects the ecological crisis (month 18) and leadership scandal (month 34) before conventional trust surveys register the decline. The early detection lead time reflects the structural sensitivity of coherence operators relative to outcome-level aggregate statistics.

as a governance coherence mechanism. See Figure 3 for the decision logic, and Table 1 for a summary of all six protocols.

5.1 Protocol 1: Collective Physiological Synchrony (Δ_1)

Hypothesis. Governance sessions preceded by 5 minutes of synchronized breathing and rhythmic coordination will produce higher trust, better decisions, and more cooperative behavior than sessions preceded by asynchronous individual activity or silence.

Design. Three-arm randomized controlled trial. Arm A: synchronized breathing exercise with rhythmic clapping (synchronous). Arm B: individual reading period with no coordination (asynchronous). Arm C: 5 minutes of silence with no instruction (control). $N \geq 36$ per arm ($N = 108$ total). Participants engage in a standardized governance simulation (budget allocation with competing priorities) after the intervention. *Power justification:* based on the meta-analytic estimate of $d = 0.58$ (95% CI: 0.40–0.77) for synchrony effects on cooperative behavior [Mogan et al., 2017], $N = 36$ per arm provides 80% power to detect $d = 0.55$ at $\alpha = 0.05$ (one-tailed) for pairwise comparisons.

Metrics. Inter-brain EEG synchrony (hyperscanning). HRV coherence (wearable monitors). Post-session trust scale (validated instrument, e.g., the Organizational Trust Inventory adapted from Mayer et al. 1995). Decision quality (rated by blinded expert panel). Cooperative surplus (economic measure from the allocation task).

Predicted outcome. Arm A shows significantly higher EEG synchrony ($p < 0.01$), HRV coherence ($p < 0.05$), trust ratings ($d > 0.5$), and cooperative surplus ($d > 0.4$) relative to Arms B and C. These effect sizes are consistent with Wiltermuth and Heath [2009] ($d = 0.48$–0.74) and the pooled meta-analytic estimate [Vicaria and Dickens, 2017].

Falsification criterion. If Arm A does not produce significantly higher trust ratings and cooperative surplus than Arm C ($p < 0.05$, one-tailed), the collective physiological synchrony

Table 1: Summary of falsification protocols. Each row specifies the operator, experimental design, total sample size (N), predicted primary effect size (d), power at $\alpha = 0.05$, and the meta-analytic source for the effect size estimate.

Operator	Design	N	Predicted d	Power	Source
Δ_1 Synchrony	3-arm RCT	108	0.55	80%	Mogan et al. [2017]
Δ_2 Narrative	A/B + follow-up	110	0.60	80%	Bruner [1991]
Δ_3 Transparency	Crossover	80	0.40	85%	Grimmelikhuijsen et al. [2013]
Δ_4 Ecological	Matched-pair quasi	12+	0.80	80%	Folke et al. [2010]
Δ_5 Restraint	2-arm RCT	70	0.50	80%	Owens et al. [2013], Fishkin [2009]
Δ_6 Environment	Within-subjects	25	0.50	80%	Kellert [2008]

operator is not supported as a governance coherence mechanism.

5.2 Protocol 2: Narrative Institutional Memory (Δ_2)

Hypothesis. Governance communications framed in narrative terms (institutional origin accounts, shared identity, purpose-driven framing) will produce higher participant retention, stronger institutional identification, and more positive sentiment than informationally equivalent communications in procedural/utilitarian framing.

Design. A/B test with longitudinal follow-up. Group A receives all institutional communications in narrative-framed versions. Group B receives informationally equivalent communications in procedural/utilitarian framing. $N \geq 55$ per group ($N = 110$ total, accounting for ∼10% attrition over 8 weeks). Duration: 8 weeks. *Power justification:* for $d = 0.60$ on the primary outcome (identity coherence), $N = 45$ completers per group provides 80% power at $\alpha = 0.05$ (two-tailed).

Metrics. Participant retention (attendance at voluntary governance events). Sentiment analysis of participant communications. Identity coherence survey (pre/post). Memory retention test (can participants articulate the institutional narrative at week 8?).

Predicted outcome. Group A shows $\geq 20\%$ higher retention, significantly stronger identity coherence ($d > 0.6$), and higher narrative memory scores ($d > 0.8$) at week 8.

Falsification criterion. If Group A does not show significantly higher retention and identity coherence than Group B ($p < 0.05$), the narrative coherence operator is not supported.

5.3 Protocol 3: Transparency (Δ_3)

Hypothesis. Governance systems with open decision dashboards (real-time visibility into decisions, rationales, and outcomes) will produce higher participation rates, higher trust scores, and lower complaint rates than systems with closed information architectures.

Design. Within-subjects crossover design. Two governance units of comparable size and function. Unit 1 operates with open dashboards for 4 weeks, then closed for 4 weeks. Unit 2 operates in the reverse order. $N \geq 40$ per unit ($N = 80$ total). The crossover controls for unit-level confounds; a 1-week washout period between conditions mitigates carryover effects. *Power justification:* the within-subjects design reduces error variance; for a crossover with expected $d = 0.40$ and within-subject correlation $r = 0.50$, $N = 40$ per unit provides 85% power at $\alpha = 0.05$. The Worldwide Governance Indicators [Kaufmann et al., 2010] provide benchmarks for "voice and accountability" scores against which the transparency manipulation can be calibrated.

Metrics. Participation rates (attendance, comment submissions, vote counts, mapped to Fung 2006's participation intensity taxonomy). Trust survey scores (administered biweekly using validated scales). Complaint rates (formal and informal). Information-seeking behavior (FOIA-equivalent requests).

Predicted outcome. Open-dashboard periods show $\geq 15\%$ higher participation, significantly higher trust scores ($d > 0.4$), and $\geq 25\%$ lower complaint rates. The trust effect may be moderated by baseline trust levels, consistent with Grimmelikhuijsen et al. [2013]'s finding that transparency effects are stronger in low-trust populations.

Falsification criterion. If open-dashboard periods do not produce significantly higher trust scores than closed periods ($p < 0.05$), the transparency operator is not supported as a coherence mechanism.

5.4 Protocol 4: Ecological Grounding (Δ_4)

Hypothesis. Governance systems that peg policy decisions to biophysical metrics (soil carbon, water quality, biodiversity indices) will produce better ecological outcomes and higher long-term institutional trust than systems that make equivalent decisions without biophysical accountability.

Design. Matched-pair quasi-experimental design. Identify ≥ 6 pairs of comparable governance jurisdictions (similar size, demographics, economic base). One member of each pair adopts ecologically-pegged governance protocols; the other maintains existing protocols. Duration: 12 months minimum (ecological metrics require at least one full seasonal cycle). *Power justification:* with 6 matched pairs and repeated ecological measurements (quarterly), a mixed-effects model with jurisdiction-level random intercepts provides 80% power to detect a standardized mean difference of $d = 0.80$ on ecological outcomes — a large effect, appropriate for a quasi-experimental design with high measurement noise. The V-Dem project's deliberative component index [Coppedge et al., 2022] provides baseline governance quality controls for each jurisdiction.

Metrics. Ecological outcomes: soil organic carbon (g/kg), water quality index, local biodiversity index (species counts, Shannon diversity). Institutional outcomes: public trust survey, voter/participant turnout, policy satisfaction ratings.

Predicted outcome. Ecologically-pegged jurisdictions show measurably better ecological metrics ($p < 0.05$ on at least 2 of 3 ecological measures) and no worse institutional trust (non-inferiority, $\delta < 0.2$ SD).

Falsification criterion. If ecologically-pegged jurisdictions do not show significantly better ecological outcomes on at least 2 of 3 measures ($p < 0.05$), or if they show significantly worse institutional trust ($p < 0.05$, inferiority test), the ecological grounding operator is not supported.

5.5 Protocol 5: Institutional Restraint (Δ_5)

Hypothesis. Governance decision-making sessions that include a structured deliberative restraint protocol (mandatory pause before final decisions, structured uncertainty acknowledgment, consultation requirements) will produce lower conflict rates, fewer decision reversals, and better physiological regulation than sessions using competitive/adversarial deliberation formats.

Design. Two-arm randomized trial. Arm A: deliberative restraint protocol (structured uncertainty acknowledgment, mandatory 2-minute reflection before votes, explicit consultation of dissenting positions). Arm B: adversarial deliberation (timed debate format with winner-take-all voting, no mandatory pause). $N \geq 35$ per arm ($N = 70$ total), with each participant completing 4 sessions to establish within-subject reliability. *Power justification:* Owens et al. [2013] reported that expressed humility predicted team engagement with $\beta = 0.22$–0.34 across samples; with 4 repeated measurements per participant, $N = 35$ per arm provides 80% power for $d = 0.50$ at $\alpha = 0.05$ using mixed-effects models. Fishkin [2009]'s deliberative polling data ($d = 0.30$–0.50 for structured vs. unstructured deliberation) further supports the expected effect magnitude.

Metrics. Conflict events (raised voices, interruptions, personal attacks — coded by blinded observers using a pre-registered codebook, inter-rater reliability $\kappa > 0.7$ required). Decision reversal rate (decisions overturned within 30 days). HRV coherence during decision-making (wearable monitors, scored as proportion of spectral power in the 0.04–0.15 Hz band). Post-session affect (PANAS; Watson et al. 1988) and satisfaction ratings.

Predicted outcome. Arm A shows $\geq 40\%$ fewer conflict events, $\geq 30\%$ fewer decision reversals, significantly higher HRV coherence during deliberation ($d > 0.5$), and higher satisfaction ratings ($d > 0.4$).

Falsification criterion. If Arm A does not produce significantly fewer conflict events and decision reversals than Arm B ($p < 0.05$), the institutional restraint operator is not supported.

5.6 Protocol 6: Deliberative Environment Quality (Δ_6)

Hypothesis. Governance sessions conducted in environments meeting established environmental design standards (natural daylight >300 lux, background noise <40 dB(A), CO_2 <1000 ppm, biophilic elements present) will produce higher productivity, lower error rates, more positive affect, and greater attention stability than sessions in standard institutional environments (fluorescent lighting, acoustic tile, no natural elements, CO_2 uncontrolled).

Design. Within-subjects alternating design. The same governance body meets in a design-standard-compliant space for alternating weeks and in a standard institutional space for the other weeks. Duration: 8 weeks (4 weeks per condition). $N \geq 25$. Environmental conditions are verified via continuous sensor monitoring in both spaces. *Power justification:* within-subjects designs with 4 observations per condition per participant and expected within-subject correlations of $r = 0.60$ provide 80% power at $N = 25$ for $d = 0.50$ at $\alpha = 0.05$. The 10–15% productivity effect from daylight access studies and the affective benefits documented in biophilic design research [Kellert, 2008] support the predicted effect sizes.

Metrics. Productivity (agenda items resolved per session, adjusted for complexity by blinded coders). Error rates (procedural errors, factual errors in decisions — coded by blinded observers). Affect ratings (PANAS administered pre- and post-session; Watson et al. 1988). Attention stability (behavioral indicators: phone checks, early departures, off-topic contributions — coded from session recordings by blinded observers).

Predicted outcome. Design-standard-compliant weeks show $\geq 15\%$ higher productivity, $\geq 20\%$ fewer errors, significantly more positive affect ($d > 0.5$ on the PANAS positive affect subscale), and $\geq 25\%$ fewer attention lapses.

Falsification criterion. If design-standard-compliant weeks do not produce significantly more positive affect and higher productivity than standard-environment weeks ($p < 0.05$), the deliberative environment quality operator is not supported.

5.7 Composite Falsification Logic

The six protocols are designed to be independently executable — each tests a single operator in isolation. However, the composite governance coherence score $G_c(t)$ makes a stronger claim: that the six operators are *jointly* predictive of governance outcomes. Testing this composite claim requires a seventh protocol.

Protocol 7 (Composite). Identify ≥ 20 governance systems of comparable size and mandate. Measure all six operators using the protocols described above. Compute $G_c(t)$ for each system. Over a 24-month observation period, track governance outcomes: institutional trust (Edelman/Pew methodology), civic participation rates, decision reversal rates, staffing stability, and ecological indicators where applicable. *Falsification criterion:* if the rank-order correlation between G_c and a composite governance outcome index is not significantly positive ($\rho > 0$, $p < 0.05$), the composite framework is not supported. If individual operators are supported but the composite is not, the multiplicative aggregation structure is wrong and alternative functional forms (e.g., additive, weighted, or threshold-based) should be investigated.

The protocols are ordered by implementation feasibility: Protocols 1, 5, and 6 can be conducted in laboratory or quasi-laboratory settings within weeks. Protocol 2 requires 8 weeks. Protocol 3 requires 9 weeks (including washout). Protocol 4 requires at least 12 months. Protocol

7 (composite) requires 24 months minimum. We recommend beginning with Protocols 1 and 6, which have the strongest existing empirical backing and the lowest resource requirements.

6 Discussion

6.1 Connection to Existing Frameworks

The Restoration Matrix is not proposed in a vacuum. It builds on and integrates several established research programs.

Ostrom [1990] identified eight design principles for successful commons governance, including clearly defined boundaries, collective-choice arrangements, monitoring, graduated sanctions, conflict-resolution mechanisms, and nested enterprises. The restoration operators provide a measurement framework for the conditions that make Ostrom's principles *operational*. Collective-choice arrangements (Δ_3: transparency, Δ_5: institutional restraint) work differently when participants are physiologically synchronized (Δ_1) and share a coherent institutional narrative (Δ_2) than when they are fragmented and alienated. Monitoring is more effective when the ecological substrate is explicitly integrated into governance (Δ_4). Conflict resolution depends on the physical and deliberative environment in which it occurs (Δ_6).

Putnam [2000] documented the decline of social capital in American civic life — the erosion of the associational networks, shared norms, and mutual trust that enable collective action. The restoration operators identify specific, measurable mechanisms through which social capital is produced and maintained. Social capital is not a mysterious emergent property; it is the aggregate result of collective physiological synchrony (Δ_1), narrative institutional memory (Δ_2), visible accountability (Δ_3), ecological grounding (Δ_4), institutional restraint (Δ_5), and deliberative environment quality (Δ_6). Each is independently measurable and independently manipulable. Woolley et al. [2010] demonstrated a "collective intelligence" factor (c) analogous to individual g, which predicted group performance across diverse tasks and was correlated with members' social sensitivity, conversational turn-taking equality, and proportion of women in the group. The c factor provides a potential convergent validity check for $G_c(t)$: governance systems scoring high on our composite should also exhibit higher collective intelligence in standardized tasks. Chan et al. [2006] developed a six-dimension framework for social cohesion measurement — social relations, task relations, perceived unity, emotions, social inclusion, and equality — that maps onto several of our operators and provides validated survey instruments for cross-validation.

The resilience literature [Folke et al., 2010, Folke, 2006, Walker and Salt, 2012, Walker et al., 2004] emphasizes the capacity of social-ecological systems to absorb disturbance and reorganize while retaining essential function. Holling [2001] formalized the adaptive cycle (exploitation, conservation, release, reorganization) and the panarchy of nested cycles that governs system dynamics across scales. The governance coherence score $G_c(t)$ provides a quantitative indicator of this capacity: a system with high G_c can absorb shocks because its structural conditions — synchrony, narrative coherence, transparency, ecological grounding, institutional restraint, deliberative environment quality — provide the redundancy and adaptive capacity that resilience requires. Governance systems in the "conservation" phase of the adaptive cycle (high connectedness, high capital) are simultaneously most efficient and most vulnerable to collapse; $G_c(t)$ is designed to detect the brittleness that conventional performance metrics miss during this phase.

The deliberative democracy tradition provides complementary normative foundations. Habermas [1996] grounds legitimate governance in communicative rationality — the quality of reasons exchanged in public discourse rather than the aggregation of pre-formed preferences. Fishkin [2009] operationalizes this through deliberative polling, demonstrating that informed deliberation produces opinion changes of $d = 0.30$–0.50 on average across hundreds of polls in 29 countries. Fung and Wright [2003] documents institutional innovations — participatory budgeting in Porto

Alegre, community policing councils in Chicago, stakeholder environmental governance — that embody high scores on Δ_3 (transparency), Δ_5 (institutional restraint), and Δ_1 (collective co-presence during deliberation). These are not utopian proposals; they are functioning institutions whose outcomes are empirically documented.

6.2 What This Framework Is Not

The Restoration Matrix does not replace formal governance structures. Laws, constitutions, regulations, electoral procedures, and judicial systems are necessary infrastructure. The claim is not that coherence is sufficient for governance — a coherent mob is still a mob. The claim is that formal structures are *not* sufficient: a governance system can have excellent formal structures and still fail if its coherence infrastructure is degraded. The six operators measure the conditions that determine whether formal structures produce their intended effects.

The framework is also not culturally prescriptive. It does not specify which rituals, narratives, or deliberative formats a governance system should adopt. It specifies that *some* mechanism must produce physiological synchrony, narrative coherence, transparency, ecological accountability, institutional restraint, and deliberative environment quality — and it provides measurement tools to assess whether the chosen mechanisms are working. A Quaker meeting and a Maori powhiri produce Δ_1 through different cultural forms; the operator measures the physiological outcome, not the cultural form.

6.3 Limitations

Several limitations constrain the current framework.

Threshold calibration. The coherence thresholds referenced from our prior work ($\Delta > 0.7$ healthy, $0.3 < \Delta < 0.7$ transition, $\Delta < 0.3$ failure), including the critical threshold $G_c^* = 0.3$ used in Proposition 1, were established on engineering data (turbofan engines, industrial valves) and require domain-specific recalibration for governance applications. We do not claim these thresholds transfer directly; they serve as initial estimates to be revised through empirical governance studies.

Equal weighting. The multiplicative structure of $G_c(t)$ implies that the six operators are equally weighted, which may not hold across all governance contexts. A governance system where ecological grounding (Δ_4) is irrelevant (e.g., a purely digital governance body) should not be penalized for low Δ_4. Weighting calibration requires empirical data from diverse institutional settings, and context-specific operator subsets may be more appropriate than the full six-operator model.

Measurement feasibility. Several measurement protocols (particularly EEG hyperscanning for Δ_1 and longitudinal ecological surveys for Δ_4) are resource-intensive and may not be feasible for routine governance assessment. We have prioritized HRV coherence as the primary physiological measure for Δ_1 precisely because of its lower deployment cost, but even wearable HRV monitoring requires participant consent and compliance that may not be achievable in all governance contexts.

Normalization and measurement invariance. The choice of normalization method (min-max vs. z-score vs. percentile rank) for mapping raw scores to $[0, 1]$ affects the composite $G_c(t)$. Min-max normalization requires established population-level ranges for each instrument, which do not yet exist for several of our measures. Cross-cultural measurement invariance is a particular concern: survey instruments validated in one cultural context may not maintain their psychometric properties in another. Grimmelikhuijsen et al. [2013] demonstrated that transparency effects on trust are moderated by cultural context, suggesting that the functional form of at least Δ_3 may require culture-specific parameterization.

Ecological validity. The falsification protocols are designed for controlled or quasi-experimental settings; real governance operates in open systems where confounds are difficult to eliminate. The quasi-experimental design for Protocol 4 (ecological grounding) partially addresses this, but the 12-month duration may be insufficient for ecological metrics to respond to governance changes.

Demand characteristics and expectancy effects. Researcher expectancy effects are a concern for any framework that measures subjective states (affect, trust, identity). Our protocols specify blinded assessment where possible, but participants in governance sessions cannot be blinded to the experimental condition (e.g., harmonically designed vs. standard environment in Protocol 6). We recommend supplementing self-report measures with behavioral and physiological indicators that are less susceptible to demand characteristics.

Operationalization of Δ_6. Deliberative environment quality remains the most difficult operator to operationalize rigorously. While sensory suitability (S_h), environmental stability (E_s), and occupant affect (A_f) have established measurement protocols, design quality (D_q) relies on post-occupancy evaluation instruments whose construct validity for governance-specific settings has not been independently established. We propose using the BUS methodology as a starting point but acknowledge that governance-specific validation studies are needed.

7 Conclusion

A governance system whose participants cannot articulate why it exists, who do not trust its processes, whose decisions ignore ecological reality, whose leaders maximize available power, and whose deliberations occur in environments that undermine cognitive function — such a system is structurally incoherent regardless of the quality of its legal framework. The gap between procedural adequacy and functional governance is not a residual to be explained away; it is the central phenomenon.

The Restoration Matrix provides a measurement framework for this structural coherence. The six operators — collective physiological synchrony (Δ_1), narrative institutional memory (Δ_2), transparency and social trust (Δ_3), ecological grounding (Δ_4), institutional restraint (Δ_5), and deliberative environment quality (Δ_6) — are measurable conditions that affect trust, participation, recovery capacity, and institutional stability. Each is grounded in existing empirical research, linked to validated measurement instruments, and — most importantly — falsifiable. Each operator comes with an explicit experimental protocol, an *a priori* power analysis, and a criterion under which it would be rejected.

Governance coherence as defined here is a composite measurement of whether the structural conditions for functional governance are present. Like the coherence score for an engine or a cardiac rhythm [Thorarinson and Hensgen, 2026], it can degrade before conventional failure metrics trigger — and like those physical systems, early detection of coherence degradation creates the possibility of intervention before failure becomes irreversible.

The framework invites empirical testing. If governance systems that score high on $G_c(t)$ do not outperform those that score low on measures of trust, participation, resilience, and ecological sustainability, the framework is wrong. That is the point. A governance theory that cannot be wrong cannot be useful.

References

David Beetham. *The Legitimation of Power.* Palgrave Macmillan, 1991.

Frank Biermann, Kenneth Abbott, Steinar Andresen, Karin Bäckstrand, Steven Bernstein, Michele M. Betsill, Harriet Bulkeley, Benjamin Cashore, Jennifer Clapp, Carl Folke, et al. Navigating the anthropocene: Improving earth system governance. *Science*, 335(6074):1306–1307, 2012.

Jerome Bruner. The narrative construction of reality. *Critical Inquiry*, 18(1):1–21, 1991.

Joseph Chan, Ho-Pong To, and Elaine Chan. Reconsidering social cohesion: Developing a definition and analytical framework for empirical research. *Social Indicators Research*, 75(2): 273–302, 2006.

Michael Coppedge, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, et al. V-Dem codebook v12. Varieties of Democracy (V-Dem) Project, University of Gothenburg, 2022.

Lee J. Cronbach and Paul E. Meehl. Construct validity in psychological tests. *Psychological Bulletin*, 52(4):281–302, 1955.

Andreas Duit and Victor Galaz. Governance and complexity—emerging issues for governance theory. *Governance*, 21(3):311–335, 2008.

Guillaume Dumas, Jacqueline Nadel, Robert Soussignan, Jacques Martinerie, and Line Garnero. Inter-brain synchronization during social interaction. *PLoS ONE*, 5(8):e12166, 2010.

Émile Durkheim. *The Elementary Forms of Religious Life*. Allen & Unwin, 1912. English translation by Carol Cosman, Oxford University Press, 2001.

Edelman Trust Institute. 2025 Edelman trust barometer. Technical report, Edelman, 2025. URL https://www.edelman.com/trust/2025/trust-barometer.

James S. Fishkin. *When the People Speak: Deliberative Democracy and Public Consultation*. Oxford University Press, 2009.

Carl Folke. Resilience: The emergence of a perspective for social–ecological systems analyses. *Global Environmental Change*, 16(3):253–267, 2006.

Carl Folke, Stephen R. Carpenter, Brian Walker, Marten Scheffer, Terry Chapin, and Johan Rockström. Resilience thinking: Integrating resilience, adaptability and transformability. *Ecology and Society*, 15(4):20, 2010.

Archon Fung. Varieties of participation in complex governance. *Public Administration Review*, 66(s1):66–75, 2006.

Archon Fung and Erik Olin Wright. *Deepening Democracy: Institutional Innovations in Empowered Participatory Governance*. Verso, 2003.

Stephan Grimmelikhuijsen, Gregory Porumbescu, Boram Hong, and Tobin Im. The effect of transparency on trust in government: A cross-national comparative experiment. *Public Administration Review*, 73(4):575–586, 2013.

Jürgen Habermas. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. MIT Press, 1996.

C. S. Holling. Understanding the complexity of economic, ecological, and social systems. *Ecosystems*, 4:390–405, 2001.

Daniel Kaufmann, Aart Kraay, and Massimo Mastruzzi. The worldwide governance indicators: Methodology and analytical issues. *World Bank Policy Research Working Paper*, (5430), 2010.

Stephen R. Kellert. Dimensions, elements, and attributes of biophilic design. *Biophilic Design: The Theory, Science, and Practice of Bringing Buildings to Life*, pages 3–19, 2008.

Roger C. Mayer, James H. Davis, and F. David Schoorman. An integrative model of organizational trust. *Academy of Management Review*, 20(3):709–734, 1995.

Rollin McCraty, Mike Atkinson, Dana Tomasino, and Raymond Trevor Bradley. The coherent heart: Heart–brain interactions, psychophysiological coherence, and the emergence of system-wide order. *Integral Review*, 5(2):10–115, 2009.

Reneeta Mogan, Ronald Fischer, and Joseph A. Bulbulia. Prosocial consequences of interpersonal synchrony: A meta-analysis. *Zeitschrift für Psychologie*, 225(3):168–186, 2017.

Elinor Ostrom. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, 1990.

Bradley P. Owens, Michael D. Johnson, and Terence R. Mitchell. Expressed humility in organizations: Implications for performance, teams, and leadership. *Organization Science*, 24(5): 1517–1538, 2013.

Richard V. Palumbo, Marisa E. Marraccini, Lisa L. Weyandt, Oliver Wilder-Smith, Heather A. McGee, Siwei Liu, and Matthew S. Goodwin. Interpersonal autonomic physiology: A systematic review of the literature. *Personality and Social Psychology Review*, 21(2):99–141, 2017.

Karl R. Popper. *The Logic of Scientific Discovery*. Hutchinson, 1959.

Robert D. Putnam. *Bowling Alone: The Collapse and Revival of American Community*. Simon and Schuster, 2000.

Roy A. Rappaport. *Ritual and Religion in the Making of Humanity*. Cambridge University Press, 1999.

Diego A. Reinero, Suzanne Dikker, and Jay J. Van Bavel. Inter-brain synchrony in teams predicts collective performance. *Social Cognitive and Affective Neuroscience*, 16(1-2):43–57, 2021.

Richard Sosis and Candace S. Alcorta. Signaling, solidarity, and the sacred: The evolution of religious behavior. *Evolutionary Anthropology*, 12(6):264–274, 2003.

Joel Thorarinson and Allison Hensgen. From prediction to discoverative intelligence: A coherence-based AI framework for detecting system drift before failure. *arXiv preprint*, 2026.

Ivana M. Vicaria and Leah Dickens. To be in synchrony or not? a meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, 72:13–20, 2017.

Björn Vickhoff, Helge Malmgren, Rickard Åström, Gunnar Nyberg, Seth-Reino Ekström, Mathias Engwall, Johan Snygg, Michael Nilsson, and Rebecka Jörnsten. Music structure determines heart rate variability of singers. *Frontiers in Psychology*, 4:334, 2013.

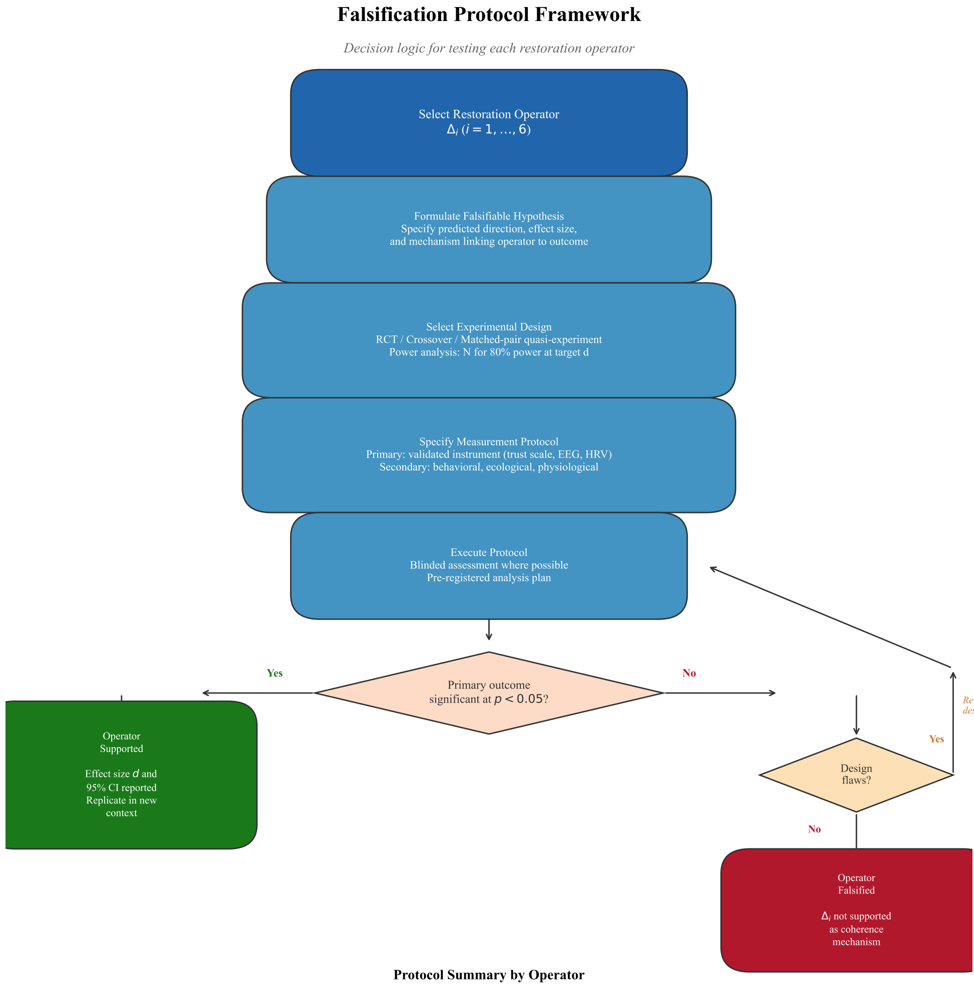
Brian Walker and David Salt. *Resilience Practice: Building Capacity to Absorb Disturbance and Maintain Function*. Island Press, 2012.

Brian Walker, C. S. Holling, Stephen R. Carpenter, and Ann Kinzig. Resilience, adaptability and transformability in social–ecological systems. *Ecology and Society*, 9(2):5, 2004.

David Watson, Lee Anna Clark, and Auke Tellegen. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6):1063–1070, 1988.

Scott S. Wiltermuth and Chip Heath. Synchrony and cooperation. *Psychological Science*, 20(1): 1–5, 2009.

Anita Williams Woolley, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone. Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330(6004):686–688, 2010.



Protocol Summary by Operator

Operator	Domain	Design	Min. Effect	Sample
Δ_1	Ritual	RCT (3-arm)	$d > 0.5$	$N = 108$
Δ_2	Narrative	A/B longitudinal	$d > 0.6$	$N = 110$
Δ_3	Transparency	Crossover	$d > 0.4$	$N = 80$
Δ_4	Ecological	Matched-pair	$d > 0.8$	6 pairs
Δ_5	Humility	RCT (2-arm)	$d > 0.5$	$N = 70$
Δ_6	Aesthetic	Within-subj.	$d > 0.5$	$N = 25$

Figure 3: Falsification protocol decision logic. For each operator Δ_i, the framework proceeds from hypothesis formulation through experimental design (with *a priori* power analysis), measurement, and execution to a binary decision: if the primary outcome is significant at $p < 0.05$, the operator is supported; if not, and no design flaws are identified, the operator is falsified. The table summarizes design parameters for all six protocols.