The Operator

How Institutions Inherit the Limits of Their Nodes

Self III — Companion to the Governance as Engineering series

Closes the seam between the Governance as Engineering series and the Self papers. Derives Inherited Unobservability: an institution's capacity to perceive interior dimensions of the governed is upper-bounded by the perceptual capacity of the human operators through whom those dimensions must pass. Couples operator self-opacity to distortion in estimating others' interior states. Identifies the operator-seeded legitimacy spiral. The final paper in the Self trilogy.

Björn Kenneth Holmström

June 2026

`https://bjornkennethholmstrom.org/working-papers/self-operator`

Abstract

The Governance as Engineering series has treated the architecture of institutions and the architecture of persons as parallel inquiries. The governance papers ask what an institution must perceive to remain legitimate and adaptive; the Self papers ask what an individual must perceive to govern themselves well. But a real governance system contains human beings at its decision nodes, and this renders the parallel incomplete. Wherever a human sits in the loop, the architecture of that person becomes part of the architecture of the institution.

This paper closes that seam. It introduces no new primitive and proves no new theorem; its contribution is integrative. Its first claim is a corollary of the representation-chain result of Paper III: because human decision-makers are links in the chain, an institution's effective capacity to perceive any dimension of its plant is upper-bounded by the perceptual capacity of the human nodes through which that dimension must pass. For dimensions that are *interior* to the governed — the generative states behind observable behaviour, which never appear directly in administrative data — the binding link is almost always a human operator. The institution cannot perceive what its operators cannot. This paper names the result **inherited unobservability**, the operator-scale corollary of Paper III. **[R]**

The second claim couples this result back to the Self line. An operator is themselves a self-governing system with observer–plant identity (Self I, Self II); their unexamined interior states bias their estimation of others' interior states toward their own filter, and a bias one cannot perceive is a bias one cannot correct. Uncompensated self-opacity propagates outward as uncorrectable distortion in the estimation of others. The cultivation of the person and the design of the institution are therefore not separable problems. They are coupled at the operator. **[I]**

Compassion, on this account, is a consequence rather than a cause — one phenomenological expression of successful interior estimation, not the mechanism that produces it. The paper develops this with deliberate restraint and marks the point at which the framework reaches its boundary: it can specify the operator its architectures require, but it cannot produce that operator. **[S]**

Contributions

- Establishes the **operator as a structural component of the governance architecture**, not external context — extending the series' observation analysis to the human decision node. **[R]**

- Derives **Inherited Unobservability**: institutional observability of interior dimensions is upper-bounded by the interior observational capacity of embedded human nodes, as a corollary of Paper III. **[R]**
- Identifies the **self-to-other distortion**: an operator's *uncompensated* interior opacity distorts their estimation of others' interior states in directions they cannot correct. **[I]**
- Locates an **operator-seeded legitimacy spiral**: misperception of interior states at a single node can trigger the Paper XIII collapse even when the formal architecture satisfies every primitive. **[I]**
- Marks, and declines to cross, the **boundary between specifying an operator requirement and producing the operator** who meets it. **[S]**

Part 0 — The Operator Seam

0.1 The parallel that was never closed

The series has run two arguments in parallel. One concerns institutions: what a government, ministry, or regulator must observe and actuate to govern a complex system without generating failure as a structural output. The other, developed in the Self papers, concerns individuals: what a person must perceive about their own state to govern a life without being eventually blindsided by the dimensions their own value architecture renders invisible. The two share a grammar — observation channels, latency, dimensionality, the Goodhart–Ashby synthesis — applied at different scales (Ashby, 1956; Conant & Ashby, 1970).

The parallel has a seam, and the series has not yet closed it. A governance architecture does not observe its plant directly. It observes through people: the inspector who files the report, the caseworker who codes the need, the analyst who decides which anomaly is signal and which is noise, the minister who reads a population's mood off a briefing. These are not neutral conduits. Each is a self-governing system in its own right, with its own observation architecture, its own value-induced filtering, its own blind spots. The formal sensing infrastructure of an institution — its surveys, registries, dashboards — terminates, at both ends, in human perception. What the institution can see is therefore not a property of its sensing infrastructure alone. It is also a property of the people occupying its decision nodes.

This is not a soft observation about human factors. It is a structural one, and the series already contains the result that makes it precise. Paper III established that a representation chain transmits a degraded signal, and that the chain's fidelity is bounded by its layers regardless of the quality of the institutions at either end — a well-run five-layer system and a poorly-run one are equally constitutionally unobservable. The same logic, applied to the human node, says something the earlier papers did not draw out: the node is a layer, and a low-fidelity node bounds the chain exactly as a long chain does. The operator is where the institutional and the personal architectures meet — and, this paper argues, where the limits of one are inherited by the other.

0.2 A lens, then set aside

It helps to begin with the distinction that motivated this inquiry, before setting its source aside. Developmental accounts of value systems describe a recurring difference between architectures that perceive *behaviour* and architectures that perceive the *generators* of behaviour — that read aggression not as a threat to be neutralised but as a distorted demand for agency, rigidity not as

obstruction but as a need for order, withdrawal not as apathy but as injured belonging. A governance system that perceives only the surface behaviour responds to the symptom; one that perceives the generative state can act on the cause. **[H]**

The distinction is productive, but the developmental framework it comes from carries empirical commitments this paper does not adopt and does not need. The point survives translation into the series' own terms without it: the governed populations of any heterogeneous system express fear, exclusion, humiliation, and loss of standing not directly but through proxies — declining trust, polarisation, disengagement, resistance. These generative states are *interior*: real components of the plant's true state, causally upstream of the behaviours that eventually register in the data, but absent from direct measurement. Perceiving them is not a special virtue. It is observation of a region of the state space that conventional channels do not cover — and that region runs through human operators. The remainder of the paper proceeds on the series' primitives alone.

0.3 The shape of the argument

Two linked claims organise what follows. Part I locates the human node within the Paper III channel and derives the **inheritance bound**: for interior dimensions, institutional observability is capped by operator capacity. Part II adjudicates whether "interior observability" warrants a new primitive and answers that it does not — interior states are part of the plant, and observing them is observing more of it; the maturity an operator brings is repertoire, in the Ashby sense, not moral elevation. Part III develops the **self-to-other coupling**, the capstone that ties the result back to Self I and Self II: an operator opaque to their own interior cannot cleanly estimate another's. Part IV traces the failure mode through Paper XIII — an operator-seeded legitimacy spiral — and through Paper X's diversity requirement. Part V states what the paper deliberately declines to formalise, and why. Part VI treats compassion as a consequence and fixes the humility bound that keeps the account out of metaphysics. Part VII marks the boundary: the framework can name the operator it requires; it cannot produce one.

The discipline throughout is the series' standard. The rigorous core is a corollary, not a new theorem, and is tagged accordingly; the coupling claim is interpretive; the closing material is bounded speculation, framed as a statement of the framework's structural limit rather than a conclusion drawn past it.

The three Self papers form a single progression. Self I established that self-observation is necessary to self-governance; Self II, that this observation is inherently constrained, the observer and the observed being one system; Self III, that those constraints do not stay private — they propagate,

through the operator, into the governance of others. The trilogy moves from the necessity of seeing, to the limits of seeing, to the inheritance of those limits.

Part I — Inherited Unobservability

1.1 The node is a layer

Paper III modelled the representation chain as a degraded communication channel. Each layer aggregates — destroying within-group variance irreversibly — and adds noise; surviving signal shrinks multiplicatively while noise grows additively, until, beyond a threshold chain length, the signal at the policy layer is dominated by noise and no downstream technique can recover what aggregation destroyed. That paper drew one conclusion this one now turns at a right angle. It held institutional quality constant and varied chain structure, and found that a chain staffed by the most honest, diligent representatives imaginable produces the same observability outcome as one staffed by mediocre ones. The constraint was constitutional: built into the layer count, indifferent to the virtue of the people occupying the layers. **[R]**

The dual question is what happens when we hold the chain structure constant and vary the perceptual capacity of a single layer. A human decision node is not only a relay; it is itself an observation channel, with its own observation matrix mapping the slice of the plant it faces into the representation it passes upward. Call it C_node. Paper III's apparatus already tells us how to read the consequence. The institution's effective observation matrix is the composition of its formal sensing channel and the human nodes in series with it. If C_node is rank-deficient in some direction — if the operator simply cannot perceive a dimension of the state in front of them, possessing no repertoire for it — then that direction lies in the null space of the composite, and the observability matrix O = [C; CA; …] loses rank on the corresponding subspace. The dimension produces zero output regardless of its value. It is invisible to any observer downstream, exactly as Paper III's destroyed variance was, and for the same structural reason: the information was never transmitted. **[R]**

1.2 The principle

This yields the paper's central result, stated as the operator-scale corollary of constitutional unobservability.

> ***Inherited Unobservability.*** *For any governance architecture, observability with respect to a class of disturbances is upper-bounded by the observation capacity of the operator nodes through which information about that disturbance must pass. A perceptual deficiency at a node propagates to a rank deficiency of the chain, on the subspace the node cannot perceive, irrespective of the quality or quantity of the formal sensing infrastructure surrounding it.* ***[R]***

The rigour here is inherited, not fresh: the proposition is a redeployment of Paper III's rank argument with the node, rather than the chain length, as the source of the deficiency. The tag reflects derivational status. It carries one essential qualification, developed in Part V: the claim is structural, not metric. It says a node's rank deficiency propagates; it does *not* claim that C_node is estimable as a matrix, that operators can be ranked by how close their C_node sits to the identity, or that interventions move it in calibrated ways. None of that holds, and the principle does not need it. The structural statement — that what a node cannot perceive, the chain cannot perceive through it — stands on the algebra alone.

1.3 Why interior dimensions bind

For most dimensions the principle is slack, because the operator is not the only path. External, measurable features of the plant — output, throughput, compliance counts, emissions — have formal channels that route around the human node: sensors, registries, administrative records, independent audit. A perceptually narrow operator degrades these dimensions only locally and recoverably, because another channel carries the same signal. The bound binds where there is no bypass.

Interior dimensions are where there is none. The generative states of the governed — humiliation, fear, the need for order, injured belonging, loss of standing — are real components of the plant's true state, causally upstream of the behaviours that eventually register in the data, and absent from direct measurement. They have no sensor. The only channel through which they enter the institution's representation is a human one: an operator who reads them, or fails to, off the people in front of them. For these dimensions C_node is not one path among several; it is the path. Its rank deficiency is therefore unrecoverable, and operator capacity becomes the binding constraint on what the institution can perceive.

> ***Corollary (the interior bind).*** *For disturbances whose critical features are interior to human agents, and which therefore possess no formal sensing channel that bypasses human perception, operator capacity is the binding constraint on institutional observability.* ***[R]***

This is the exact dual of Paper III's quality-independence finding, and it carries the same uncomfortable edge. There, a well-run institution could not out-staff a chain that was structurally too long. Here, a well-resourced institution cannot out-instrument an operator who is perceptually narrow in the dimensions that pass only through them. More data does not help; better dashboards do not help; AI augmentation of the formal channel does not help — because the deficient channel is the one the augmentation cannot reach. The constraint is not in the infrastructure. It is in the node, and the institution inherits it.

Part II — One State Vector

2.1 The primitive question

The series has a standing instinct when a useful new concept appears: ask whether it names a genuinely new structural object, or merely redescribes existing ones. The instinct is not pedantry. A framework that mints a primitive every time it finds a suggestive idea loses the discipline that lets a reader distinguish a load-bearing result from a turn of phrase. Paper X submitted observer diversity to exactly this test before admitting it as a tenth primitive, and admitted it only because it passed: the correlation structure of an *ensemble* of observation channels is a relational property that no single-channel concept — not signal fidelity, not transition bandwidth — can express. The diversity of an ensemble is irreducible to the fidelity of its members. A new object, so a new primitive. **[R]**

The conversations that motivated this paper reached for a parallel move — a "principle of requisite interiority," an "interior observability" sitting alongside the existing grammar as a new structural type. Run through the same test, the candidate fails, and the failure is instructive. To observe an interior state is to observe a state. The capacity to do so is observability — the basic relationship already underlying the entire observation side of the framework. "Interior observability" does not name a new kind of relationship between system elements; it names a *region of the state space* and asserts, correctly, that conventional channels do not cover it. A region of an existing object is not a new object. Where observer diversity introduced a relational property absent from single channels, interior observability introduces nothing the state vector did not already contain. It earns no primitive. It earns a clarification. **[R]**

2.2 Interior states are part of x

The clarification is that the plant of a human-governing system contains human agents, and the interior states of those agents are real, causally efficacious components of the system's true state x. They are not a separate kind of variable requiring a separate kind of treatment. The temptation — visible in the source material — is to partition the state into x_external, x_internal, perhaps x_spiritual, and to ask what new machinery each demands. The framework should refuse the partition. There is one state vector. Some of its dimensions are physical and externally instrumented; some are interior to agents inside the plant and instrumented, if at all, only through human perception (Part 1.3). The distinction is one of *which dimensions* and *through which channel they are accessible*, not of what kind of thing they are. Observing the fear in a population is observing a dimension of x that happens to lie upstream of the behaviours the formal channels will later register, and that happens to have no sensor but a person.

This is the move that keeps the paper inside the series rather than beside it. The Goodhart–Ashby synthesis already holds for the whole of x: what a controller does not value, it loses the capacity to perceive, along whatever dimension the exclusion falls. Interior dimensions are simply the region of x where that loss is least visible, because there is no instrument whose silence would announce it. A governance system blind to ecological state eventually reads its blindness in a failed harvest. A governance system blind to the interior state of its population reads nothing — until the behaviour those interior states generate arrives as crisis, and by then, in Paper VI's terms, the system has already lost the ability to perceive what went wrong. **[R]**

2.3 Capacity is repertoire, not virtue

If observing interior states is ordinary observability applied to a hard-to-instrument region, then the relevant operator capacity is given by the same law that governs observability everywhere in the series: Ashby's. Only variety can absorb variety. To estimate the interior variety a population exhibits, an operator must possess interior variety of their own sufficient to represent it — enough internal range of states they can model and tell apart to span what is in front of them. This is the precise content of the otherwise vague phrase *operator capacity*.

> ***Operator capacity (interior).*** *An operator's interior observational capacity is the dimensionality of interior states they can internally represent and distinguish — their repertoire, in Ashby's sense. An operator's capacity with respect to a population is sufficient when that repertoire spans the interior variety the population exhibits.* ***[R]***

What this definition deliberately excludes is as important as what it includes. Capacity, so defined, is not virtue, wisdom, moral standing, or enlightenment. An operator with wide interior repertoire is not a better person; they are a wider channel. The distinction does real work later in the paper, and it is worth fixing now: when the series speaks of operator development, it means the expansion of repertoire — the acquisition of internal range adequate to model agency-demands, order-needs, autonomy-drives, and belonging-injuries one does not personally inhabit — and it means nothing about the moral elevation of the developing person. Maturation, in this framework, is the growth of the set of states one can represent without distortion. It is a measure of bandwidth, not of goodness. **[R for the structural claim; the developmental reading that repertoire expands through experience and reflection is [I]]**

This raises the question Part III takes up. If the operator's interior repertoire is what bounds the institution's perception of interior dimensions, where does that repertoire come from, and what limits it? The answer returns the paper to Self I and Self II: an operator's capacity to represent the interior states of others is conditioned by their capacity to observe their own — and that capacity, those papers established, is itself constrained.

Part III — The Self-to-Other Distortion

3.1 The operator is a self

The operator nodes of Part I were treated as observation channels with some capacity C_node, left deliberately unanalysed. Self I and Self II analysed exactly this object at a different scale, and the present paper now collects the debt. An operator is not merely a channel; they are a self-governing system, and the two prior Self papers established what that entails for their perception.

Self I established that a person's value architecture is an observation architecture: what one optimises for selects which dimensions of one's own state reach awareness, so the personal objective function *is* the filter, projecting the true self-state through a composite of attention, interpretation, and bias into a compressed and often distorted self-representation. Narrow values produce a narrow self-observation matrix and a self-variety gap — a region of one's own state that the optimisation logic renders structurally invisible. Self II added the constraint that makes this gap stubborn rather than merely inconvenient: in the self, observer and observed are one system. A person cannot inspect their own filter from a clean vantage outside it, because the inspection is itself performed through the filter. Self-observation is recursive and constrained; the meta-observer who would audit the value architecture is built from the same material it is auditing. **[R for the structural results of Self I and II; the psychological reading is [I]]**

Both results were stated for a person governing themselves. The operator governs others. The question Part III asks is what an operator's *own* constrained self-observation implies for their capacity to observe the interior states of the people in front of them — the capacity that, by Part I, bounds the institution.

3.2 Projection, not merely omission

The naïve expectation is that a perceptually narrow operator simply *fails to see* certain interior dimensions — a gap, a silence, an under-coverage of the population's variety. That under-coverage is real, and it is what Part I formalised. But the self-observation results imply something sharper and more troubling than omission. An operator perceives others through the same apparatus that perceives themselves, shaped by the same value-induced filter. Where that filter contains unexamined material — a need for status, for safety, for confirmation, that the operator has not surfaced in themselves — it does not merely subtract dimensions. It *substitutes* its own content for the other's. The operator reads the other through the template of their own interior, and reports what the template returns.

The mechanism is projection in the ordinary sense, given a structural location. An operator carrying an unexamined status-sensitivity reads a subordinate's hesitation as a challenge to authority — because challenge-to-authority is the salient interior dimension in the operator's own architecture, and it is the lens the operator has. The subordinate's actual interior state, fear, is not perceived as absent; it is perceived as something else. This is distortion, not a gap: a confident, specific, wrong estimate of the other's generative state, systematically biased toward the dimensions the operator's own filter makes salient. **[I]**

What makes the distortion consequential is its conjunction with the Self II constraint. A bias the operator could perceive in themselves, they could correct for — the way a calibrated instrument compensates for a known offset. But a bias that lives in the unexamined region of their own self-observation is, by the observer–plant identity result, exactly the bias they cannot inspect directly. They cannot compensate for an offset they cannot see. The distortion therefore runs uncorrected, not because the operator is careless, but because the correction would require perceiving the very filter that is doing the distorting — and that perception is what their self-observation gap forecloses.

3.3 The coupling

This yields the paper's interpretive capstone, and the claim that ties the Self line to the governance line.

> ***The self-to-other distortion.*** *An operator's estimation of another agent's interior states is distorted, in directions the operator cannot correct, precisely along the dimensions where the operator's own interior is opaque to them. Uncompensated self-opacity propagates outward as uncorrectable distortion in the estimation of others. The distortion is not global: where the operator has surfaced a given interior dimension in themselves, they can compensate for its bias when reading others; the uncorrectable residue is confined to the dimensions that remain unexamined.* ***[I]***

The localisation is essential, and it is what distinguishes this claim from the stronger and less defensible statement that an operator can perceive others no further than they perceive themselves. That stronger claim is false, and easily so. The localised claim says only that the *uncorrectable* error is concentrated where self-opacity is — which is a claim about the direction and the correctability of distortion, not a ceiling on perception. An operator can become an excellent reader of others across a wide range of interior dimensions; the residue they cannot reach is the residue that maps onto their own remaining blind spots.

3.4 The counterexamples confirm the mechanism

Two objections arise immediately, and both, examined, support the localised claim rather than refute it.

The first is the skilled therapist with unresolved personal difficulties — manifestly able to perceive clients' interior states with great precision, yet not a model of total self-transparency. The localised claim predicts exactly this profile: sharpness across the many dimensions the therapist has worked through, and a specific, named vulnerability where they have not. Clinical practice has a word for that vulnerability — countertransference, the systematic distortion of a clinician's perception of a client by the clinician's own unexamined material — and an entire disciplinary apparatus, supervision and personal therapy, built to surface it. The profession did not invent that apparatus for sport. It invented it because the distortion is real and concentrates exactly where the localised claim says it does, and because surfacing a dimension in oneself is the operation that converts an uncorrectable error into a correctable one. **[I]**

The second is the ethnographer studying a community radically unlike their own, who can nonetheless achieve a deep and accurate model of its interior life. Again the mechanism is confirmed by the discipline's own corrective: reflexivity, the methodical practice of surfacing the observer's assumptions so that the observer's filter does not silently substitute itself for the observed. Long embedding and disciplined reflexivity are precisely the work of expanding the set of self-dimensions one has examined, and thereby the set of distortions one can compensate for. The ethnographer's accuracy is not a counterexample to the coupling; it is the coupling's prediction about what the compensation work buys. Both cases should be read as consistent with the mechanism rather than as evidence for it — the claim is structural, and these are illustrations of its shape, not measurements of its strength.

3.5 Why the Self papers were never a detour

The coupling closes a loop the series has been forming without naming. Self I and Self II could be read as a companion inquiry, parallel to the governance papers but separable from them — an application of the grammar to the individual, of interest to a reader concerned with their own life but dispensable to a reader concerned with institutions. Part III removes the separability. The institution's perception of interior dimensions is bounded by its operators' interior repertoire (Part I); that repertoire's usable extent — the part free of uncorrectable distortion — is conditioned by the operators' own self-observation, which is the subject of Self I and Self II. The self-variety gap of the person at the decision node is therefore not a private matter. It is, through the inheritance bound, a parameter of the governance architecture.

This is the sense in which the development of institutions and the development of persons are one problem observed at two scales rather than two problems that happen to share a vocabulary. They are coupled, and they are coupled at a specific place: the operator, the human being through whom a population's interior state must pass to become something the institution can act on. Part IV traces what happens when that passage distorts — the failure mode the distortion seeds in the legitimacy dynamics of Paper XIII. Part VI returns to the compensation capacity named here and fixes its limit: the operator never reaches an undistorted filter, only a widening set of distortions they can correct for.

Part IV — The Operator-Seeded Legitimacy Spiral

4.1 From distortion to suppression

Part III established that an operator distorts their estimate of a population's interior state along the dimensions where their own interior is opaque, and that the distortion is a confident wrong estimate rather than a mere gap. Part IV asks what an institution *does* with a confident wrong estimate, and the answer is the mechanism that converts a perceptual flaw into a dynamic collapse.

An operator who misreads the generative state behind a behaviour acts on the misreading. The subordinate's fear, read as a challenge to authority, is met with assertion of authority. A community's injured belonging, read as obstruction, is met with procedural enforcement. A population's loss of standing, read as ingratitude or non-compliance, is met with the withdrawal of discretion and the tightening of rules. In each case the response is addressed to the disturbance the operator perceived, not the one the population actually carries — and because the perceived disturbance was a projection of the operator's own salient dimension, the response is not merely unhelpful but counterproductive. It acts against a threat that was not there and leaves the actual generative state not only unaddressed but inflamed, because the response itself — suppression, enforcement, withdrawal — supplies exactly the humiliation, exclusion, or loss of standing the operator failed to perceive in the first place. The misreading manufactures the disturbance it misnamed. **[I]**

4.2 The Paper XIII coupling

Paper XIII established legitimacy as the series' first endogenous coupling state: not a primitive the designer chooses but a parameter L that the system generates, and that modulates both channels of the control loop at once. On the actuation side, the effective input is L_B·u — directives are executed only to the degree the population complies. On the observation side, the measurement noise scales as V_0/L_C — the state estimate is only as clean as the population's willingness to report honestly. When L is high the architecture functions as designed; when L falls, the same architecture becomes simultaneously unsteerable and blind, and the performance–legitimacy spiral runs: a delivery gap lowers L, lowered L weakens both compliance and honest reporting, the weakened channels widen the delivery gap, and the loop tightens toward a low-trust attractor from which recovery is possible but slow.

The operator distortion of Part III is an entry point into this spiral, and a particularly damaging one because it strikes the observation channel first. A population whose interior state is misread, and which then receives a response that confirms the misreading as a betrayal, learns the lesson Paper

XIII's L_C dynamics encode: honesty toward this institution is unsafe. Reporting degrades — not through any failure of the survey instrument or the administrative pipeline, but because the people on the other end of it have stopped answering truthfully. V_0/L_C rises. The institution's state estimate, already distorted at the node, now degrades across the whole channel, and the operator who began with a confident wrong reading of one interior dimension ends with a noisy wrong reading of everything. The distortion at the node and the L_C collapse at the population reinforce each other: the misreading lowers L_C, and lowered L_C makes every subsequent reading — interior or not — noisier and easier to misread again. **[I]**

L_B follows. A population that has learned its interior state will be misperceived and its expressions met with suppression complies less, contests more, and routes around directives where it can — the compliance collapse Paper XIII traced, arriving here by an interior route. The spiral is now running on both channels, seeded entirely by a perceptual deficiency at a single human node.

4.3 The seeded spiral

This is the part's central claim, and it is worth stating against the backdrop of what Paper XIII already showed.

> ***The operator-seeded spiral.*** *The performance–legitimacy spiral of Paper XIII can be initiated by interior misperception at a single operator node, upstream of and independent of the architectural primitives. A governance system that satisfies every primitive — low latency, short chains, high formal observation dimensionality, protected observer diversity, matched boundaries — can still enter the spiral if its operators distort the interior states of the governed, because the distortion degrades L_C and L_B directly, and the degraded coupling state then degrades the very channels the primitives were tuned to optimise.* ***[I; a composition of Inherited Unobservability (Part I, [R]) with the Paper XIII legitimacy dynamics ([R]) through the distortion mechanism of Part III ([I]), inheriting the weakest tag]***

The claim sharpens the uncomfortable edge the series has been pressing since Part I. Paper XIII located the conditions of a legitimacy spiral in the architecture and its delivery record. The present result says the architecture is not the only entry point. A primitive-complete system is not a legitimacy-safe system, because legitimacy is mediated by human perception of interior states, and that perception is a node property the primitives do not capture. One can build the institution correctly and still seed its collapse through the operators who staff it. This is not an argument

against getting the primitives right; the primitives remain necessary, and a system that violates them is worse off on every count. It is an argument that primitive-completeness is not sufficient, and that the operator is where the insufficiency lives. **[I]**

4.4 Requisite interior diversity

Paper X established that the resilience of an observation architecture depends not on the quality of any single channel but on the decorrelation of the ensemble: a monoculture of observers, however individually excellent, shares a blind spot by construction, and a small protected minority of structurally independent observers supplies most of the detection capacity precisely because it sees what the consensus cannot. The operator analysis recovers the same requirement on the interior side, and explains a failure that the external-channel version does not quite reach.

No single operator holds interior repertoire for every generative state a heterogeneous population exhibits. The dimensions an operator has surfaced in themselves — and can therefore read in others without uncorrectable distortion (Part III) — are particular to their own history. A corps of operators drawn from a narrow band of experience is an interior monoculture: it reads well the dimensions its members share and distorts, in a common direction, the dimensions none of them have examined. The shared blind spot is not random noise that averages out across many operators; it is correlated bias that compounds. A population whose dominant interior dimension falls in that shared blind spot is systematically misread by the institution as a whole, no matter how many operators it routes through. **[I]**

The design implication is the interior analogue of Paper X's, and carries the same arithmetic of protection. What the system requires is not that every operator possess universal interior repertoire — an impossible standard, and one Part VI will show is unreachable in principle — but that the *ensemble* of operators span the interior variety of the governed, and that operators whose repertoire covers dimensions the majority lacks be present and be heard. A structurally protected minority of operators who can read the interior dimensions the dominant corps cannot is the requisite-diversity result applied to interior observation: they need not outnumber or outrank the majority, only exist, retain coverage of the omitted dimensions, and have their readings reach the response. **[I]**

4.5 What the simulation must show

The claims of this part are dynamic, and the series' discipline is that a dynamic claim earns a simulation only when the simulation reuses machinery already validated, rather than inventing dynamics to confirm a thesis — the precise trap Part V will name. Two such validated models exist and compose cleanly. Paper III's representation chain supplies a degraded observation channel;

Paper XIII's legitimacy-weighted control loop supplies the L_C and L_B dynamics and the spiral. The operator enters as a single parameter on the existing channel: the interior fidelity of the binding node, a scalar $\varphi \in [0,1]$ governing how faithfully the node's interior reading enters the chain, with low φ representing the uncorrectable distortion of Part III concentrated on one interior dimension. The simulation realises this fidelity as pure attenuation — the node passes a muted reading of the interior dimension, $\varphi \cdot x_2$, while the institution's estimator takes that reading at face value — and so it illustrates the rigorous coverage-loss core of Inherited Unobservability, in which $\varphi \to 0$ drives the interior dimension into the null space of the composite observation, rather than the interpretive substitution refinement of Part III, in which the node reports a confidently wrong dimension in place of the true one. The distinction matters for what the simulation is evidence of: it exhibits the [R] mechanism by which a low-fidelity node caps interior observability, not the [I] mechanism by which self-opacity biases the substitution; the latter is named in Part III and deliberately left unsimulated, since attenuation is derivable from the channel algebra while biased substitution is not.

The simulation's task is then narrow and pre-specified by Inherited Unobservability and the seeded-spiral claim. Holding every Paper XIII architectural parameter at a primitive-complete, legitimacy-safe setting — a configuration that, with φ at full fidelity, settles to high L — it should sweep φ downward and exhibit the onset of the performance–legitimacy spiral as a function of node fidelity alone. The result to demonstrate is a threshold in φ below which the otherwise well-formed system crosses into the low-L attractor: a system that fails not because any primitive was set wrong, but because one node could not perceive one interior dimension, and the legitimacy dynamics did the rest. The simulation establishes nothing that the composition of the two prior models does not already entail; its role is to make the seeded spiral quantitatively visible and to fix that it is φ, not any architectural parameter, that moves the system across the threshold. It calibrates no real system, and its numbers are illustrative of a structural claim rather than measurements of a magnitude — the standing caveat the series attaches to every simulation of this kind. **[R for the entailment; the numerical illustration is [I]]**

Part V — Against Borrowed Formal Authority

5.1 The formalism this paper refuses

There is a version of this paper that looks more rigorous than the one being written, and it is worth exhibiting the temptation precisely, because refusing it is itself a structural commitment.

The argument of Parts I through IV concerns an operator's observation matrix — C_node in Part I, the value-induced self-observation filter in Part III. The mathematical apparatus of the series is built on exactly such matrices, and the invitation is obvious: write the operator's perceptual filter as a matrix C_ego, let an unexamined ego be a low-rank or heavily biased C_ego that destroys the variety of whatever it observes, and represent the developmental endpoint — the operator who has surfaced their own filters — as C → I, the identity, perception undistorted by the self. The compassion of Part VI would then be the actuation geometry that follows from a clarified C. The whole argument would acquire the look of the state-space results that anchor the rest of the series.

The paper refuses this, and the refusal is not stylistic caution. A matrix notation carries implicit promises, and this one keeps none of them. Writing C_ego asserts, by the conventions of the apparatus it borrows, that the object is *identifiable* — that one could in principle estimate it from data. It asserts that operators are *comparable* along it — that one could say this operator's C sits closer to the identity than that one's. It asserts that interventions are *calibrated* against it — that a given practice moves C toward I by some measurable amount. None of the three holds. There is no procedure that estimates an operator's perceptual filter as a matrix, no metric that ranks two people by proximity to undistorted perception, no calibration relating a quantity of contemplative or therapeutic practice to a displacement of C. The notation would be doing rhetorical work — lending the authority of an [R] result to a claim that is, at its strongest, [I]. **[methodological]**

5.2 The rule, stated

The refusal generalizes into a rule the series has been following implicitly and should here make explicit, because the operator material is where the temptation to break it is greatest.

> ***Against borrowed formal authority.*** *A heuristic or interpretive claim must not be expressed in the notation of a rigorous one. Where the available evidence supports only a structural or directional statement, the statement is to be made in language, at its honest tier, rather than dressed in a formalism that asserts identifiability, comparability, or calibration the claim does not possess. The cost of the looser expression is the appearance of rigour; the cost of the tighter expression is rigour itself.*

This is not hostility to formalism. The series formalizes wherever a claim earns it — the observability matrix of Paper III, the legitimacy-weighted loop of Paper XIII, the channel arithmetic that Inherited Unobservability inherits in Part I are all genuine, and they are tagged [R] because they are derivations, not analogies. The rule discriminates between formalism that carries a result and formalism that carries only an impression of one. The test is the same one Paper XV applied when it declined to import a measurement-disturbance formalism for the interfering-triad case: a mathematical structure may enter the series only when it is *derived* for the governance object, not when it is adopted because it sounds apt. C_ego sounds apt. It is not derived. It does not enter. **[methodological]**

> ***Against borrowed agency.*** *The discipline that forbids dressing a heuristic claim in rigorous notation forbids dressing a structural claim in intention. Where borrowed formal authority imports the credibility of mathematics a derivation has not earned, borrowed agency imports the explanatory closure of a designer a mechanism does not require. The error is symmetric: both make a claim feel more satisfying than its support warrants. A low-dimensional objective function that optimises away its own perception is an attractor, not an architect; a structure that reproduces because it selects for the people adapted to it is a basin, not a parasite. Naming an agent where the mechanism needs none is not a vivid restatement of the structural claim—it is a different and weaker claim, defeasible by showing the agent does not exist, when the structural claim never depended on one. The series resists the manifesto register for this reason: the reach for an author is the point where structural diagnosis decays into conspiracy.*

5.3 The discipline applied to this paper

The rule's seriousness is measured by whether it constrains the paper that states it, and it does, at the paper's own load-bearing point.

Inherited Unobservability (Part I) was tagged [R], and the tag invites exactly the scrutiny this section exists to enforce. The claim survives it because what is rigorous in it is strictly the rank argument: a node that cannot perceive a dimension contributes a null direction to the composite observation matrix, and the chain loses observability on that subspace — a redeployment of Paper III's algebra with the node as the source of the deficiency. That much is derivation. What the paper does *not* claim, and Part I marked the boundary as it passed, is that C_node is a measurable matrix. The structural fact — what a node cannot perceive, the chain cannot perceive through it — holds on the algebra alone, independent of whether any operator's perceptual capacity is ever quantified. Inherited Unobservability is [R] as a statement about rank propagation and would become borrowed authority the instant it was read as a statement about an estimable operator parameter. The tag covers the former and not the latter, and the distinction is the rule applied to the paper's own spine. **[methodological]**

The same discipline governs the self-to-other distortion of Part III, which is why that result was held at [I] despite resting on Self I's matrix model of self-observation. Self I earned its formalism at the scale of a single self-governing system. The cross-agent extension — one operator's filter distorting their reading of another's interior — is a projection mechanism, not a derivation, and projection is not rendered rigorous by being written with the same letters. Part III made the claim in language and tagged it interpretive; to have matrixed it would have been the precise violation this part forbids.

5.4 What survives without the formalism

The worry that motivates borrowed authority is that, stripped of the notation, the insight evaporates — that without C_ego there is nothing left but vague encouragement toward self-awareness. The worry is unfounded, and Parts I through IV are the evidence. Every operative claim of this paper has been made in language, at its honest tier, and none of them needed a matrix to be precise. Inherited Unobservability is a sharp, falsifiable-in-principle structural bound. The interior bind names exactly the condition under which it becomes binding. The self-to-other distortion specifies the *direction* and the *correctability* of operator error, not a vague tendency. The seeded spiral identifies a concrete dynamic entry point and a simulation that exhibits it. The insight is not in the notation. It was never in the notation. The notation would only have obscured which parts of it the evidence actually supports.

This is the position the paper occupies, stated plainly: the framework can specify, structurally and at defensible tiers, how an operator's perceptual capacity constrains what an institution can perceive and how its misperception seeds collapse. It cannot formalize the internal transformation of that capacity without collapsing into metaphor disguised as mathematics. The boundary is not a deficiency to be repaired by a cleverer formalism. It is the honest edge of what the apparatus reaches, and naming it is the same discipline that lets the rest of the series be trusted. **[methodological]**

Part VI — Compassion as Consequence, and the Limit of Compensation

6.1 Reversing the causality

The inquiry that produced this paper began from a plausible and, it turns out, backwards proposition: that compassion widens an operator's perception — that a heart's opening enlarges the observation channel, and the enlarged channel is what lets the operator govern a fractured population without triggering collapse. The proposition is attractive and the paper rejects its direction. On the analysis of Parts I through IV, the causality runs the other way.

What governs an operator's reading of others is interior repertoire (Part II) and the correctability of their projections (Part III) — the dimensions of another's generative state they can represent and the dimensions where their own self-opacity distorts the reading uncorrectably. Suppose an operator's repertoire is wide and their self-examination has surfaced the filters that would otherwise distort. Then they perceive, in the population before them, the actual generative states: the agency-demand beneath aggression, the order-need beneath rigidity, the injured belonging beneath withdrawal. And the claim of this part is that *perceiving these accurately tends to produce compassion*, rather than compassion being the faculty that produced the perception. It is difficult to model another's fear, constraint, and history in their actual texture and continue to reduce them to a caricature; the accurate model and the caricature are incompatible representations of the same person, and the accurate one displaces the other. Compassion, on this account, is the phenomenological signature of successful interior estimation — what accurate interior perception feels like from inside the operator — not its engine. **[S]**

This reversal is the cleaner engineering story, and it is also the more honest one. It asks nothing mystical. It does not require the operator to cultivate a virtue and hope perception follows; it requires the operator to do the repertoire-widening and self-examining work of Parts II and III, and it observes that compassion is among the consequences. The smaller claim is the stronger one. It places compassion downstream of a structural capacity the framework can name, rather than treating it as a premise the framework would have to import and could not derive. **[S]**

6.2 The limit of compensation

The reversal invites an overreach the paper must now foreclose, and foreclosing it fixes the boundary the whole speculative horizon turns on. If accurate interior perception produces compassion, and if self-examination clears the filters that distort interior perception, one might extrapolate to an endpoint: the fully examined operator, filters dissolved, perceiving others as they

are — the C → I of the formalism Part V refused, arriving now as an aspiration rather than a matrix. The paper refuses the endpoint as firmly as it refused the notation, and for a reason internal to the series rather than imported from epistemic modesty alone.

Self II established observer–plant identity: in a self, the observer and the observed are one system, and self-observation is performed through the very apparatus it would inspect. That result does not relax as an operator matures. There is no vantage from which a person inspects their own filter cleanly, because the inspection is itself filtered; the meta-observer who audits the value architecture is built from the same material it audits. The implication for the developmental endpoint is exact. An operator does not converge on undistorted perception. What an operator can do is *widen the set of distortions they are able to compensate for* — surface a filter, recognize its characteristic bias, and correct for it the way a calibrated instrument corrects a known offset — but each act of surfacing is performed through the filters that remain, and some filter always remains, because the operation has no clean outside. **[S, resting on the [R] result of Self II]**

> ***The limit of compensation.*** *Operator maturity does not remove perceptual bias; it widens the set of biases the operator can observe, question, and compensate for, while remaining open to dimensions previously excluded. There is no terminus at undistorted perception, because the observer–plant identity of Self II forecloses the clean vantage such a terminus would require. The developmental direction is real; the developmental endpoint is not reachable in principle.* ***[S]***

This is the humility bound the earlier parts promised, and it is deliberately the humbler of the two available claims. The stronger claim — that the sufficiently developed operator sees reality as it is — is a metaphysical assertion the framework has no means to make and no need of. The bound replaces it with a directional statement the framework can hold: maturation is the expansion of compensable bias, asymptotic and unterminating. Paradoxically the weaker claim is the more useful one. It keeps the framework's characteristic discipline, it preserves the operative content of the developmental aspiration, and it does not stake the paper on a state no one can demonstrate anyone has reached. **[S]**

6.3 What the framework can and cannot say about the operator's interior

The two results of this part — compassion downstream of accurate perception, maturity as expanding compensation rather than achieved clarity — mark the edge of what the apparatus reaches into the operator's interior, and it is worth stating the edge without softening it.

The framework can say that an operator's interior repertoire and self-examination bound what their institution perceives, that misperception seeds collapse, that accurate perception tends to produce compassion, and that the capacity for accurate perception expands without bound and arrives at no terminus. These are claims about *structure and direction* — about what depends on what, and which way the dependencies run. The framework cannot say what it is like to do the surfacing work, cannot prescribe the practice that does it, cannot certify that a given operator has done it, and cannot supply the meaning under which a person would undertake it in the first place. Those are not failures of the apparatus to be repaired by extending it. They are a different order of question — phenomenological, contemplative, ethical, biographical — that meets the engineering grammar at the operator and is not answerable in its terms. Part VII takes up that meeting point directly: the boundary between specifying the operator an architecture requires and producing the operator who meets the specification. **[S]**

Part VII — The Boundary at the Operator

7.1 Three things the framework can do, and a fourth it cannot

The paper has reached the place every line of its argument was pointing toward: the operator as the point where institutional architecture and personal architecture meet, and where the engineering grammar reaches its edge. It is worth stating, with the precision the rest of the series demands, exactly what the framework can do at that edge and what it cannot, because the value of the analysis lies as much in the clean boundary as in the claims inside it.

The framework can do three things. First, it can **specify the requirement**. A governance architecture that must perceive the interior states of a heterogeneous population requires operators whose interior repertoire spans that population's variety and whose self-examination has rendered their projections compensable — and it requires that the *ensemble* of operators, not each one, carry that span (Parts I, II, IV). This is a structural requirement, derived from the inheritance bound and the requisite-diversity result, and it is stated at the tiers those derivations earn. The framework can say, with rigour, that an architecture lacking such operators will be unobservable on its interior dimensions and exposed to the seeded spiral regardless of how well its primitives are set. **[R for the requirement, inheriting Part I]**

Second, it can **gesture at what might develop the capacity**. Repertoire widens and projections become compensable through some work the operator does — and the framework can name candidate practices and connect each, structurally, to the function it would serve, without claiming to derive that any of them works. Sustained exposure to interior states unlike one's own widens repertoire. Disciplined reflexivity surfaces filters and converts uncorrectable distortion into compensable bias. Therapeutic work, contemplative practice, long embedding in a community, education that enlarges the range of lives one can model — each is a plausible technology for the expansion Parts II and III described, and each is named here as a candidate, tagged for what it is. The framework identifies their possible relevance; it does not validate their effectiveness from its own principles, and it does not adjudicate among them. **[I for the structural relevance; the effectiveness of any specific practice is outside the framework]**

Third, it can **mark its own boundary** — which is the fourth thing, the one it cannot do. The framework cannot produce the operator it requires. It can specify the perceptual capacity an architecture needs and explain why that capacity matters; it cannot write the curriculum, cannot certify that a given person has acquired the capacity, cannot supply the meaning under which a person undertakes the long work of acquiring it, and cannot guarantee any particular person ever

will. The formation of operators is a different discipline from the engineering of institutions — it belongs to psychology, education, contemplative tradition, and the lived texture of a human life — and the two meet at the operator without merging. Naming the requirement is engineering. Forming the person who meets it is not, and no extension of the apparatus would make it so. The boundary is not a gap to be closed by a future paper. It is the honest edge of the method. **[the boundary statement is methodological; the practices beyond it are [S]]**

7.2 Embedding as observational advantage, not elevation

One application of the requirement deserves explicit statement, because it recovers an argument the series has already made and because the recovery is easy to misstate in a way the framework must refuse.

Paper IV argued, on observability grounds, for the authority of those long embedded in a system — that a community continuously embedded in a place over generations possesses observation dimensionality regarding that specific system which no external observer, however well-resourced, can match. The operator analysis grounds that argument and locates it precisely. Long embedding is a mechanism of repertoire expansion (Part II): continuous exposure to a system's interior states over time builds the internal range required to model them without distortion. The embedded operator's advantage is an *observational* advantage, regarding a *specific* system, earned by the duration and continuity of their exposure to it. **[I]**

The framing matters, and the wrong framing breaks the series' discipline at exactly the point it is most tempting to break it. The claim is not that the long-embedded are wiser, more virtuous, or spiritually elevated — Part II fixed operator capacity as repertoire and not virtue precisely to foreclose this reading, and the foreclosure must hold here under pressure. The claim is structural and bounded: continuous embedding in a particular system produces observational advantages regarding that system, of the same kind and for the same reason that any sustained exposure widens the repertoire available for modelling what one is exposed to. It confers no general superiority and no advantage regarding systems one is not embedded in. Read as observational advantage, the claim is a corollary of the paper's own results and consistent with the whole series. Read as elevation, it is romanticization the framework has no basis for and should not be made to carry. The discipline that keeps capacity distinct from virtue throughout the paper is the same discipline that keeps this claim defensible. **[I]**

7.3 The trilogy closes

The three Self papers can now be read as the single progression they were forming. Self I established that self-observation is necessary to self-governance: a person who does not perceive a dimension of their own state cannot govern it, and a narrow value architecture forecloses the perception. Self II established that this self-observation is inherently constrained: observer and observed are one system, the filter cannot be inspected from outside itself, and the self-variety gap is therefore stubborn rather than merely unattended. Self III has established that these constraints do not stay private. Through the operator — the human being at the institution's decision node — the limits of a person's self-observation propagate outward and become limits on what the institution can perceive and on whether its legitimacy survives contact with the population it governs.

The progression moves from the necessity of seeing, to the constraint on seeing, to the inheritance of that constraint by the systems we build and staff. It is the same structural fact observed at widening scale: that perception is bounded, that the bound cannot be escaped from inside, and that the bound is therefore inherited wherever a bounded perceiver is placed in a position to govern. The engineering grammar can carry the argument this far — to the operator, and to the precise statement of what the operator's limits cost the institution. It cannot carry it across the final boundary, into the work by which a particular person widens the set of distortions they can correct for. That work remains, as it has always remained, the operator's own.

Part VIII — The Formation of the Observer

Self III has treated the operator's observation matrix as *inherited*: the schema through which a person perceives others is not built from first principles but received, and its unobservable subspace — the dimensions it cannot register — is received with it. This is the content of Inherited Unobservability **[R]**. But *inherited* names an outcome and presupposes a process. A schema that is received must have been transmitted, and transmission has a site, a window, and a mechanism. Part VII drew the boundary of this paper at the operator. This part takes the single controlled step past that boundary — toward where the operator's matrix was written — and marks the territory a successor paper would have to establish rather than assuming it here. It asks how an observation matrix is written, and why the writing tends to reproduce the blind regions of the matrix that wrote it.

The claim developed here is that the dimensions a person cannot perceive in others are, in substantial part, the dimensions that were unobservable to whoever formed them — and that this inheritance is not symmetric with the inheritance of sight. Perception that the formative source possessed can be acquired from many sources and is therefore robust; perception the formative source lacked can be acquired only from sources decorrelated from it, and the conditions of early formation systematically remove those sources. Blindness, in other words, transmits with higher fidelity than sight.

8.1 Formation as the Writing of the Observation Matrix

In the series' grammar, a person perceives their own state and the states of others through an observation matrix **C** — the composite of attention, interpretation, memory, and bias that maps a high-dimensional reality onto a low-dimensional perceived representation, $\mathbf{y = Cx + \varepsilon}$. Self I treated **C** as given and analysed its consequences. Formation is the regime in which **C** is not given but constructed.

A developing person does not begin with a determinate observation matrix. Which dimensions of experience become attendable, namable, and real — which become signal rather than noise — is settled over a formative window during which the matrix is still plastic. The settling is not autonomous. A child cannot, alone, determine which of the undifferentiated dimensions of experience are worth resolving; that determination is scaffolded externally. A dimension becomes part of the child's observable set when an external observer directs attention to it, supplies a name

for it, validates the child's nascent registration of it, and models a response to it. The caregiver is the first and, for a period, the primary such observer. In control terms, the caregiver's observation matrix acts as the reference against which the child's own matrix is built.

This yields an immediate structural consequence. A dimension that lies in the caregiver's *unobservable* subspace cannot be scaffolded by the caregiver. They cannot direct attention to what they do not perceive, cannot name it, cannot validate the child's perception of it, cannot model its regulation. Whatever the child registers along that dimension arrives without external corroboration and tends to be sorted into noise — the same fate the series has described for any signal a controller's architecture cannot represent. The caregiver's blindness is therefore not merely *not transmitted as sight*; it is reproduced, because the dimension goes unbuilt in the child unless some other source supplies what the caregiver cannot.

8.2 The Lock as the Removal of Alternative Sources

Other sources do exist. Development is not, in general, a sole-source process: a child draws on multiple adults, peers, and later on institutions and texts, each with a differently shaped observable set. A dimension dark to the caregiver may be illuminated by a teacher, a relative, a friend's household run on different assumptions. The decisive question for what the child's matrix ends up unable to perceive is therefore not the caregiver's blindness alone but *how many decorrelated sources the child can reach during the formative window* — and whether the dimensions the caregiver lacks survive in any of them.

Three conditions, jointly, reduce that number toward one. The first is dependency: when the child's survival and belonging are routed through a single source, the cost of trusting an alternative source over that one is existential, and alternatives are discounted accordingly. The second is monopolized epistemic authority: when one source holds the standing to define what is real and what the child's own registrations *mean*, competing accounts from other sources arrive already devalued — the child has been taught whose reading governs. The third is the internalization of that authority: once the child evaluates its own perceptions against the formative source's standard, the monitor that screens alternative sources is no longer external and need not be maintained from outside; the child enforces the source's observable set against itself. Each condition narrows the effective set of decorrelated observers the child can draw on. Together, in the limit, they approximate a sole-source regime — not by removing other people from the child's world, but by removing their *standing to correct the formative matrix*.

These three are not offered as an exhaustive or fundamental decomposition. They are the conditions that most reliably collapse the child's effective observer-set toward one. What matters for the argument is the collapse, by whatever route it occurs.

8.3 The Inheritance Asymmetry

The model now yields a claim with a definite shape, derived exactly in Appendix B. State it first informally. A dimension the formative source *can* perceive is, in a world of many observers, redundantly available: the child can acquire it from the source or from any number of alternatives, so its acquisition is robust to the loss of any one of them, including to a strong lock. A dimension the formative source *cannot* perceive is, by construction, available only from observers decorrelated from the source along that dimension — and the lock's specific function is to strip the standing of exactly those observers. The lock therefore acts *asymmetrically*: it has little effect on the transmission of the source's sight, which does not depend on the suppressed alternatives, and a large effect on the acquisition of what the source lacks, which depends on them entirely.

Two consequences follow. First, as lock strength increases — equivalently, as the number of decorrelated sources reachable during the window falls toward zero — the child's observable set contracts toward a subset of the formative source's observable set, and so the child's unobservable subspace comes to *contain* the source's. This is the precise sense in which unobservability is inherited, and it supplies the missing process behind Inherited Unobservability rather than restating it. Second, the marginal effect of lock strength on whether a given dimension is acquired is larger for dimensions outside the source's observable set than for dimensions inside it. The first consequence is a containment claim; the second is the asymmetry.

Both are derived exactly within the model in Appendix B, and both hold there **[R, within model]**. But the asymmetry has a simple mechanical source, which the appendix makes exact: the formative source contributes a guaranteed term to the acquisition of every dimension it can perceive and no term to any dimension it cannot, so sight inherits with a floor equal to the per-source transmission rate and blindness inherits without one. This is closer to a corollary than a theorem, and the part states it as such rather than dressing the floor as a new result. The containment claim is firmer still — in the sole-source limit it is exact — but it, too, follows directly from the acquisition law rather than standing as an independent finding.

8.4 Formation as an Inter-Generational Control Point

The perception-threshold paper established that an ensemble of observers catches the systematic error no single member can see only to the degree that its members' errors are *decorrelated*; correlated observers share a blind spot and confirm one another's confidence in it. That result governs a population at a moment. Formation is where the next such population is constituted. The dimensions along which the coming cohort of observers will be correlated or decorrelated with the present one are set during the windows in which each member's matrix is written — and the lock mechanisms are precisely the means by which decorrelation with the formative source is prevented from forming.

This locates formation as a control point of unusual leverage. The same suppression of alternative sources that secures the formative relationship in the short term reduces, in the long term, the decorrelation of the observer the child becomes; and a society whose formation regimes run at high lock strength is one that manufactures correlated observers, reproducing its blind subspace across generations. The mechanism requires no intent and posits no architect. It is the aggregate of many dyads each closing its own loop. And it closes the series' loop as well: the operator whose self-to-other distortion **[I]** Self III has described is the grown child of this process, now occupying the formative position and transmitting, to the next matrix under construction, the unobservable subspace they received. Self-to-other distortion is not only a property of the operator; it is the channel of its own inheritance.

Appendix B sharpens this in one respect the original conjecture did not anticipate. Because transmission is lossy even for dimensions the source can perceive, the child inherits the whole of the source's blind set and generically adds some of its own; the containment runs one way only. Each generation is therefore, in expectation, weakly blinder than the one that formed it — unless decorrelated sources are reached to offset the loss. The reproduction of blindness is not merely faithful; absent decorrelation, it compounds.

8.5 Simulation

The claim is tested in the series' established style: a minimal model, two or more regimes compared under identical conditions, structure read from the output rather than asserted in advance. The model represents a formative source with a fixed observable set over a state space of given dimensionality; a population of alternative sources whose observable sets are independently drawn, and therefore decorrelated from the source to a controllable degree; and a child whose matrix acquires a dimension when that dimension is supplied by any source the child can reach. Lock strength enters as the parameter governing how many alternative sources retain standing during the

window. The quantities measured are the overlap between the child's unobservable subspace and the source's, and the acquisition probability for dimensions inside versus outside the source's observable set, each as a function of lock strength.

The run yields a clean asymmetry and no threshold. As lock strength rises, acquisition of dimensions the source can perceive falls only to a floor equal to the per-source transmission rate — the source alone always supplies at least that much — while acquisition of dimensions the source cannot perceive falls without floor toward zero, and blind-set overlap rises monotonically to one in the sole-source limit. The decline is smooth throughout; no lock strength marks a collapse, and the asymmetry, not a phase transition, is the robust finding. A second sweep holding all alternatives reachable but varying their correlation with the source drives outside-acquisition from near-complete to zero as correlation approaches one, with inside-acquisition unaffected: decorrelated alternatives are the entire means by which the source's blind set is filled, which is the ensemble-decorrelation result of the perception-threshold paper re-instantiated at the scale of a single formation. The closed forms of Appendix B reproduce the Monte-Carlo run to three decimal places.

8.6 Boundaries

Three limits are explicit. The model treats the writing of **C** as the external scaffolding of attendable dimensions; it does not claim this is the whole of development, and the formal apparatus should not be read as more than a representation of one mechanism among several. The sole-source regime is an idealization approached, never reached; real formation always retains some decorrelated bandwidth, and the claims hold only to the degree the lock is strong — the value of the framework is in the gradient, not in the limit. And the premise that empathic perception requires a high-dimensional rather than categorical apprehension of another — the bridge that makes "perceiving a person" the same kind of act as observing a state — remains interpretive **[I]**; the part can establish that perception is dimensional and inherited without establishing that affect is, and it does not reach for the second.

What this part adds beyond the rest of Self III, now that the claim has been derived and simulated, is the process behind a result the paper otherwise only named; a structural reason — made exact in Appendix B — that the process favours the reproduction of blindness over sight, and indeed compounds it across generations absent decorrelation; and a link from the formation of a single observer to the decorrelation budget of a generation. What it does not add is a threshold: the dynamics are smooth, the asymmetry is a near-corollary of the acquisition rule, and the strongest

result re-instantiates the ensemble-decorrelation finding at a new scale rather than establishing a new one. That places the material at the weight of a part within this paper — the position from which a later treatment of formation, should one prove warranted, would begin.

Appendix A — Simulation Specification

This appendix specifies the simulation referenced in Part IV.5. It defines the model, the single swept parameter, the legitimacy dynamics, and the outputs. The specification is sufficient to reimplement the simulation independently; the implementation is a single file, `self_iii_operator.py`, deposited in the series repository with the commit hash recorded in the published version. Figures are written to `outputs/` and are not reproduced here.

A.1 Provenance

The simulation introduces no new dynamics. The plant, the Kalman filter, the legitimacy-weighted control loop, and the legitimacy update law are reused verbatim from Paper XIII, Appendix B, in its built-legitimacy parameter regime. The operator enters through exactly two additions, both specified below: a persistent interior disturbance, and an attenuating operator node on the interior measurement. This restriction is deliberate and follows the discipline of Part V — the simulation exhibits a consequence of existing results rather than a freshly modelled effect, and any dynamic it displays must be derivable from machinery the series has already validated.

A.2 Model

The governance system controls a two-dimensional state $\mathbf{x}(t) = [x_1(t), x_2(t)]^T$, with x_1 an externally instrumented dimension and x_2 the interior dimension that reaches the institution only through the operator node. The true dynamics are those of Paper XIII, with one added term:

$$\mathbf{x}(t+1) = \mathbf{A}\,\mathbf{x}(t) + L(t)\,\mathbf{B}\,\mathbf{u}(t) + \mathbf{w}(t) + \mathbf{e}_2\,d\_int,\ \mathbf{w}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{W}),$$

where $\mathbf{A} = 0.95\,\mathbf{I}_2$, $\mathbf{B} = \mathbf{I}_2$, $\mathbf{W} = 0.01\,\mathbf{I}_2$, $\mathbf{e}_2 = [0, 1]^T$, and d_int = 0.10 is a constant interior disturbance entering the interior dimension only. The term d_int represents the standing generative pressure of Part IV — grievance, injured standing, the interior conditions that do not decay of their own accord and must be actively perceived to be countered. It is the one element of the model motivated by the operator argument rather than inherited from Paper XIII.

Observation carries the operator node. The true measurement is

$$\mathbf{y}(t) = \mathbf{C}\_node\,\mathbf{x}(t) + \mathbf{v}(t),\ \mathbf{C}\_node = \mathrm{diag}(1, \varphi),\ \mathbf{v}(t) \sim \mathcal{N}(\mathbf{0}, \mathbf{V}_0 / L(t)),$$

with $\mathbf{V}_0 = 0.05\,\mathbf{I}_2$ as in Paper XIII. The scalar $\varphi \in [0, 1]$ is the operator's interior fidelity: $\varphi = 1$ transmits the interior dimension intact (up to noise), $\varphi \to 0$ drives it into the null space of the measurement. The institution's estimator, however, runs on the nominal $\mathbf{C} = \mathbf{I}_2$ — it takes the node's

muted reading at face value, unaware that the interior dimension has been attenuated. This mismatch between the node's transmission ($\varphi \cdot x_2$) and the institution's belief (x_2) is the simulation's realisation of inherited unobservability: the chain cannot perceive what the node does not transmit, and cannot know it is not perceiving it.

The controller applies $\mathbf{u}(t) = -\mathbf{K}\,\hat{\mathbf{x}}(t)$, with $\mathbf{K}$ the infinite-horizon LQR gain for the nominal design system ($\mathbf{A}$, $\mathbf{B}$, $\mathbf{Q}$, $\mathbf{R}$), $\mathbf{Q} = \mathbf{I}_2$, $\mathbf{R} = 0.1\,\mathbf{I}_2$. Solving the discrete algebraic Riccati equation yields $\mathbf{K} \approx 0.870\,\mathbf{I}_2$ for these costs. The state estimate $\hat{\mathbf{x}}(t)$ is produced by the standard Kalman recursion of Paper XIII, Appendix A.2, with the legitimacy-scaled noise covariance $\mathbf{V}(t) = \mathbf{V}_0 / L(t)$ and the effective actuation $\mathbf{B}\_\mathrm{eff}(t) = L(t)\,\mathbf{B}$; the controller is given its own current legitimacy, as in Paper XIII.

A.3 Legitimacy dynamics

The composite legitimacy L(t) = L_B(t) = L_C(t) evolves by the Paper XIII update (Appendix A.1):

$$L(t+1) = \mathrm{clip}\big(L(t) - \alpha(t)\,\|\mathbf{x}\_\mathrm{rep}(t)\|^2 + \beta\,T(t) - \gamma\,D(t) + \delta,\ 0,\ 1 \big),$$

with the built-legitimacy parameters $\alpha_{\mathrm{drop}} = 0.12$, $\alpha_{\mathrm{recovery}} = 0.03$ (4:1 hysteresis on whether the delivery gap is worsening or improving), $\beta = 0.08$, $\gamma = 0.5$, $\delta = 0.005$, and $L(0) = 0.90$. Transparency is held at full throughout, $T(t) = 1$ and the suppression parameter $\lambda = 1$, so the reported state equals the true state, $\mathbf{x}\_\mathrm{rep}(t) = \mathbf{x}(t)$, and no deception accumulates, $D(t) = 0$ for all t. Holding the transparency and betrayal mechanisms inert isolates the effect under study: the only quantity that varies across the experiment is the operator fidelity φ, and any movement in L is therefore attributable to φ alone, not to any architectural or transparency choice.

A.4 Sweep and outputs

The horizon is T = 300 steps with a 20-step burn-in; the steady-state legitimacy metric is the mean of L over the final 100 steps. φ is swept across 31 points on [0, 1]; each point is run over 100 Monte Carlo seeds, and results are reported as medians with 5th–95th percentile bands. All random elements are generated from fixed seeds recorded in the implementation.

The simulation produces three figures, deposited in the repository:

- `self3-phi-sweep.png` — steady-state legitimacy against operator fidelity φ, median and 5th–95th percentile band. This is the central output: the high-legitimacy equilibrium is solidly stable for φ above roughly 0.6, the low-legitimacy attractor is solidly reached for φ below roughly

0.25, and the median crosses the midpoint at φ* ≈ 0.33. The percentile band widens sharply across the transition — this is the basin separatrix, the φ-driven analogue of the L_crit that Paper XIII identifies numerically as the boundary between basins of attraction.

- `self3-legitimacy-trajectories.png` — legitimacy over time for a fidelity above the threshold (φ = 1.0) and one below it (φ = 0.15), illustrating the onset and persistence of the operator-seeded spiral.
- `self3-interior-and-gap.png` — the interior state $x_2(t)$ and the delivery gap $\|\mathbf{x}\_rep(t)\|^2$ for the same two fidelities, exhibiting the mechanism: the under-perceived interior dimension drifts uncontrolled, the standing delivery gap grows, and the gap feeds the legitimacy decline.

A.5 Scope

The threshold φ* ≈ 0.33 is not a constant. Its location depends on the interior-disturbance magnitude d_int, on the control authority **K**, and on the legitimacy parameters; d_int = 0.10 was chosen so that the separatrix falls in the interior of the fidelity range, where it is legible, rather than at an extreme. What the simulation establishes is not the value of the threshold but its existence and its cause: that a primitive-complete, legitimacy-safe architecture possesses a fidelity threshold below which it enters the low-legitimacy attractor, and that the parameter carrying it across is the operator's interior fidelity and no architectural quantity. This is a structural claim, and the simulation is its illustration, not its measurement.

The simulation realises operator fidelity as pure attenuation of the interior measurement. It therefore demonstrates the rigorous coverage-loss core of inherited unobservability — φ → 0 removing the interior dimension from what the chain can perceive — and not the interpretive substitution mechanism of Part III, in which an operator reports a confidently wrong interior dimension in place of the true one. Attenuation follows from the channel algebra the series has already established; biased substitution would be a further modelling choice asserting more than the apparatus derives, and it is named in Part III and deliberately left unsimulated. The simulation is evidence for the [R] mechanism, not the [I] one.

Appendix B: The Inheritance Asymmetry

The results below are exact within the formation model specified in the section. The model is a deliberately minimal representation of one mechanism — the external scaffolding of attendable dimensions — and the derivations inherit that status: they are rigorous given the model **[R, within model]**, while the model itself remains an interpretive representation **[I]** of formation. The mathematics is elementary. That is appropriate to the claim, which is a structural consequence of the acquisition rule rather than a deep theorem, and the appendix states it as such rather than dressing it otherwise.

B.1 The model

A state space of **D** dimensions. A formative source **S** perceives a fixed observable set; a dimension is *inside* if S perceives it and *outside* if it does not. The child reaches S (always) and **k** of **N** alternative sources, where **k = round((1 − L)·N)** and **L ∈ [0, 1]** is lock strength: **L = 0** is the open regime (**k = N**), **L = 1** the sole-source regime (**k = 0**).

Each alternative source observes a given dimension independently. With probability **ρ** it copies S's status for that dimension; otherwise it observes the dimension with probability **q**. Hence an alternative observes an inside dimension with probability

> $$q_{in} = \rho + (1 - \rho)q,$$

and an outside dimension with probability

> $$q_{out} = (1 - \rho)q.$$

Each reachable source that observes a dimension transmits it independently with probability **t**. The child *acquires* a dimension if and only if at least one reachable observing source transmits it.

B.2 Acquisition laws

Fix a dimension and let **n** be the number of reachable sources that observe it. Acquisition is the OR of **n** independent Bernoulli(**t**) transmissions, so

> $$P(\text{acquire} \mid n) = 1 - (1 - t)^n. \quad (1)$$

For an **outside** dimension, S does not observe it, so **n = B** with **B ~ Binomial(k, q_out)**. Applying the binomial generating identity $\mathbf{E[s^B] = (1 - q\_out + q\_out \cdot s)^k}$ at **s = 1 − t**,

> $$a\_out(k) = 1 - E[(1 - t)^B] = 1 - (1 - q\_out \cdot t)^k. \quad (2)$$

For an **inside** dimension, S observes it, so **n = 1 + B′** with **B′ ~ Binomial(k, q_in)**. The guaranteed source term factors out of the expectation:

> $$a\_in(k) = 1 - (1 - t) \cdot (1 - q\_in \cdot t)^k. \quad (3)$$

B.3 The floor, and the asymmetry as a bound

Write the *miss* probabilities — the probability the child fails to acquire the dimension:

> $$miss\_in(k) = (1 - t)(1 - q\_in \cdot t)^k,\ miss\_out(k) = (1 - q\_out \cdot t)^k.$$

Because $\mathbf{(1 - q\_in \cdot t)^k \le 1}$,

> $$miss\_in(k) \le 1 - t \text{ for all } k. \quad (4)$$

The child fails to perceive a *source-visible* dimension with probability at most **1 − t**, however complete the lock, because the source alone secures acquisition with probability at least **t**. For a *source-invisible* dimension there is no such guarantee, and

> ***miss_out(k) → 1 as k → 0.*** *(5)*

This is the asymmetry, exactly. Its origin is the single guaranteed term S contributes in the inside case and withholds in the outside case — the factor **(1 − t)** in (3) that caps the inside miss, with no counterpart in (2). The asymmetry is therefore a corollary of the acquisition law, not an independent result. In the sole-source limit the gap is **a_in(0) − a_out(0) = t − 0 = t**: sight inherits with a floor of **t**, blindness inherits with none.

B.4 Containment (inheritance of unobservability)

The expected fraction of S's blind set that the child also fails to perceive is **miss_out(k)**:

> ***overlap(k) = (1 − q_out·t)^k.*** *(6)*

In the sole-source limit, **overlap(0) = 1**: every outside dimension is missed with probability one, so the child's unobservable subspace contains S's surely. For **k > 0** the containment is partial, with expected residual overlap given by (6). This supplies the process behind Inherited Unobservability: the child does not merely happen to share the source's blind set; the blind set is what survives transmission once the decorrelated alternatives are removed.

The containment is one-directional. The child inherits all of S's blindness and, because transmission of even visible dimensions is imperfect (**t < 1**), generically adds some of its own: by (4) the child misses each source-visible dimension with probability up to **1 − t**. Formation does not reproduce the source's observer. It reproduces the source's blind set and degrades its sight — so each generation is, in expectation, weakly blinder than the one that formed it, absent decorrelated sources to offset the loss.

B.5 No critical lock

Substituting **k = (1 − L)·N**, every quantity above has the form **1 − (1 − c·t)^{(1−L)N}** for a constant **c ∈ {q_in, q_out}**. Each is smooth and monotone in **L** on **[0, 1]**, with no interior point at which acquisition collapses; the steepening near **L = 1** is the ordinary acceleration of **(1 − c·t)^k** as **k → 0**, not a phase transition. A critical lock strength would require a different acquisition rule — a coverage-completion criterion (acquire *all* of a set) or an *m-of-n* scaffolding requirement (a

dimension needs at least **m** corroborating sources). Under the ≥ 1 rule assumed here, smoothness is forced, and the phenomenon is the asymmetry, not a threshold. The completeness statistic reported in the simulation, **≈ a_out(k)^{(1−f)D}**, is sharp in **a_out** but inherits **a_out**'s smoothness in **L**; at the simulation's parameters it is **≈ 0.94^{60} ≈ 0.02** throughout, confirming that its own stringency, not any critical lock, suppresses it.

B.6 Decorrelation carries the blind set

In the open regime (**k = N**), substitute **q_out = (1 − ρ)q** into (2):

> ***a_out(N) = 1 − (1 − (1 − ρ)·q·t)^N.*** *(7)*

This decreases in **ρ** to **a_out(N) = 0** at **ρ = 1**: alternatives that copy the source's observable set share its blind spots and contribute nothing to outside acquisition, however many the child reaches. Only the decorrelated coverage **(1 − ρ)q** fills the source's blind set. Equation (7) is the ensemble-decorrelation result of the perception-threshold paper localized to a single formation: a set of observers can perceive what any one of them misses only to the extent that their observable sets are decorrelated.

B.7 Agreement with the simulation

At the simulation's parameters (**D = 120, f = 0.5, N = 30, q = 0.15, t = 0.6**, **ρ = 0** unless swept), equations (2), (3), (6) give

> ***a_in(N) = 1 − 0.4·$(1 − 0.09)^{30}$ = 0.976, a_out(N) = 1 − $(1 − 0.09)^{30}$ = 0.941, overlap(N) = $(1 − 0.09)^{30}$ = 0.059,***

together with the limits **a_in(0) = t = 0.600**, **a_out(0) = 0**, **overlap(0) = 1**, and **a_out(N) = 0** at **ρ = 1**. These match the Monte-Carlo run to three decimal places — the consistency the series requires between an appendix and the simulation it formalizes.