From Minimal Overminds to Self-Grounding Epistemic Monads

Generativity, Abstraction, Meta-Learning, and Reflective Epistemic Revision in Interacting Husserlian Monadology

Abstract

This paper develops a sequence of increasingly powerful finite and countably infinite intentional systems within the framework of interacting Husserlian monadology introduced in *Possible Worlds of Interacting Husserlian Monads* [1]. The companion technical note *Minimal Finite Models of Diachronic and Self-Conscious Overminds* [2] established, under explicit binary-independence assumptions, that a strongly diachronic overmind requires at least eight global states, that indexical self-consciousness need not enlarge this state space, and that independently variable reflective self-attention requires sixteen states.

Starting from these results, we ask a succession of minimality questions: What is the smallest overmind that can imagine a genuine counterfactual? What is the smallest that can invent an intentional object not previously available to it? What architecture permits indefinitely many inventions? What additional structure is required for abstraction, recombination, learning rules, learning rules for producing rules, revising the criteria by which patterns are discovered, and finally revising the standards by which epistemic revisions themselves are evaluated?

A sixteen-state model suffices for an independently counterfactually imaginative overmind possessing a fixed counterfactual noema. A twenty-four-state architecture suffices for a finite inventor that passes through three distinct statuses—unavailable, imagined, and willed—for a newly constructed noema. Permanent open-ended invention cannot be implemented by any fixed finite state space if indefinitely many novel noemata must remain internally distinguishable. The minimal state-space cardinality therefore jumps to $\aleph_0$. A free unary term algebra on one primitive noema provides a minimal model of open-ended linear generativity.

Richer cognition is then classified not by state-space cardinality but by the type and arity of its generative operations. A free binary term algebra permits open-ended recombination. Contrastive abstraction of two co-accessible noemata yields reusable one-hole schemas, producing a minimal abstractive-recombinational monad. Comparison of learned schemas in turn yields higher-order meta-schemas—rules for transforming rules. We next make the abstraction policy itself mutable, obtaining a self-revising epistemic monad whose criterion of structural similarity can be altered by experience. Finally, the evaluator used to compare epistemic policies is itself placed inside the revisable constitution. Infinite regress is avoided by a model of diachronic reflective equilibrium in which transitions are supported by locally shared justificatory bridges between predecessor and successor epistemologies.

The resulting investigation suggests a candidate minimal invariant kernel for reflective epistemic agency: reflective self-addressability, nondegenerate coupling to givenness, and diachronic bridgeability. Classical logic, probability, simplicity principles, ontology, theories of truth, and even standards of epistemic

evaluation can in principle be moved into the revisable part of the monad. The progression therefore reveals a qualitative hierarchy not merely of increasingly many mental states but of increasingly deep forms of self-transformation.

---

1. Introduction

The earlier paper *Possible Worlds of Interacting Husserlian Monads* [1] proposed a research program rather than a single psychophysical theory. Fundamental reality was provisionally represented as an ecology of monadic processes whose intrinsic character is panprotopsychist, while sufficiently developed monads possess a broadly Husserlian organization involving retention, primal presentation, protention, intentionality, horizons, synthesis, and an ego-pole.

A central methodological decision was to postpone the question of which such structures are realized in our world. Instead, one studies a space of possible psychophysical worlds and asks:

$$\boxed{\text{What forms of subjecthood and intentional interaction are formally coherent?}}$$

This led naturally to toy worlds containing telepathic copying, phenomenal sharing, fusion, fission, overminds, nested societies of subjects, collective intention, and collective imagination.

The subsequent technical note *Minimal Finite Models of Diachronic and Self-Conscious Overminds* [2] changed the style of the project. Rather than merely displaying possible architectures, it began asking **minimality questions**.

Under a simple binary architecture it distinguished:

$$4$$

states for a weak relational overmind;

$$8$$

states for a strongly diachronic overmind carrying collective memory not recoverable from the instantaneous states of its constituents;

the same

$$8$$

dynamical states for an indexically self-conscious overmind when the self-noema is treated as a rigid intentional index; and

$$16$$

states when explicit reflective self-attention must vary independently.

The present paper continues that program.

The central question becomes:

$$\boxed{\text{What is the least structure required for progressively stronger forms of cognitive and epistemic self-transformation}}$$

The resulting sequence moves through:

$$\text{imagination,}$$

$$\text{invention,}$$

$$\text{open-ended generativity,}$$

$$\text{abstraction and recombination,}$$

$$\text{meta-rule learning,}$$

$$\text{revision of inductive bias,}$$

and finally

$$\text{revision of epistemic standards themselves.}$$

A striking change occurs along the way. For the smallest minds, the natural complexity measure is the number of states:

$$4, \quad 8, \quad 16, \quad 24.$$

Once open-ended invention is demanded, finite state cardinality becomes impossible. The state space becomes countably infinite, and thereafter the meaningful measure of cognitive complexity is no longer cardinality but **generative signature, type hierarchy, memory architecture, and depth of revisability**.

The project thus evolves from a finite phenomenological complexity theory into something closer to a theory of **self-transforming intentional processes**.

---

2. Background: the eight-state diachronic overmind

We begin with the principal construction of [2].

Let two persistent constituent subjects have binary local states

$$a, b \in \{0, 1\}.$$

Define the collective relation

$$d = a \oplus b.$$

Interpret:

$$d = 0$$

as a higher-order noema J, representing jointness or intentional agreement, and

$$d = 1$$

as D, representing divergence.

A weak overmind could consist entirely of d. But then

$$d_t = f(a_t, b_t)$$

at every moment, so its state synchronically supervenes on the lower pair.

The stronger model introduces a binary collective memory

$$m.$$

One explicit reversible dynamics is

$$\boxed{a' = a, \qquad b' = a \oplus m, \qquad m' = a \oplus b.}$$

Since

$$d = a \oplus b,$$

the induced collective dynamics is

$$\boxed{(d, m) \mapsto (m, d).}$$

Thus:

$$m_{t+1} = d_t$$

and

$$d_{t+1} = m_t.$$

A minimal Husserlian interpretation is:

$$R_C(t) = R_{m_t},$$

$$I_C(t) = R_{d_t},$$

$$P_C(t) = R^*_{m_t}.$$

The collective past is retained, the current relation is presented, and the retained relation determines the protended next relation.

Most importantly, two global states can have identical instantaneous constituent states but different memories:

$$(a, b, m) = (0, 0, 0)$$

and

$$(0, 0, 1).$$

They yield different futures.

Therefore:

$$(a_t, b_t)$$

alone is insufficient for a deterministic Markovian description.

The higher collective state is synchronically irreducible while remaining diachronically generated:

$$m_t = d_{t-1}.$$

This motivated the distinction:

$$\boxed{\text{synchronic supervenience } /} \qquad = \text{diachronic}$$

Under the binary independence assumptions stated in [2], eight states are both necessary and sufficient.

---

3. Self-consciousness and the distinction between semantic and dynamical cost

The eight-state overmind can be equipped with a persistent ego-pole

$$E_C$$

and a rigid self-noema

$$\sigma_C$$

such that

$$\rho(\sigma_C) = E_C.$$

The self-noema does not itself constitute an independently varying dynamical variable.

It functions as an intentional index:

$$\mathsf{Self}_C(q) = \langle \sigma_C, q \rangle.$$

Thus C can represent:

$$\text{“we are joint”},$$

$$\text{“we were divergent”},$$

or

$$\text{“we shall be joint again”}$$

without enlarging the underlying eight-state dynamics.

This produced an important methodological distinction:

$$\boxed{\text{semantic complexity need not equal dynamical state complexity.}}$$

If, however, C must be able independently to switch between ordinary intentionality and explicit self-thematization,

$$s = 0$$

versus

$$s = 1,$$

then a new binary degree of freedom appears. Under an independence assumption,

$$8 \times 2 = 16.$$

This result becomes a template for the later imagination theorem.

---

4. The smallest independently imaginative overmind

Protention is not yet imagination.

A protention represents something like:

> this is the future toward which the current intentional process is tending.

Imagination requires an additional modal separation:

> this could occur although it is not what the present dynamics currently produces.

Let the ordinary eight-state overmind obey:

$$d' = m.$$

Introduce a counterfactual noema:

$$\boxed{Z: \qquad d' = \neg m.}$$

Its content is approximately:

> instead of restoring the retained collective relation, instantiate its opposite.

Suppose Z is already available in C's noematic repertoire. Give C a binary noetic mode:

$$q = 0: \quad \text{Imagine}(Z),$$

$$q = 1: \quad \text{Will}(Z).$$

Define:

$$\boxed{a' = a,}$$

$$\boxed{b' = a \oplus m \oplus q,}$$

$$\boxed{m' = a \oplus b,}$$

together, for example, with

$$q' = 1 - q.$$

Then:

$$d' = m \oplus q.$$

Hence:

$$q = 0 \Rightarrow d' = m,$$

while

$$q = 1 \Rightarrow d' = \neg m.$$

The same noema Z is first entertained under imagination and subsequently adopted as a project whose adoption makes a causal difference.

Under the requirement that both attitudes be independently available over every underlying eight-state configuration:

$$|S| \geq 8 \times 2 = 16.$$

The explicit construction attains the bound.

Theorem 1. Minimal independently counterfactual imagination

Within the binary-independent architecture of [2], an overmind possessing a fixed collective noema Z, and able independently to distinguish

$$\mathrm{Imagine}(Z)$$

from

$$\mathrm{Will}(Z)$$

while allowing this distinction to affect subsequent dynamics, requires at least sixteen global states. Sixteen suffice.

The conceptual novelty is the appearance of:

$$\boxed{\text{modal surplus.}}$$

The mind now represents not only what is, was, or is expected, but a structured possibility that the operative dynamics does not presently realize.

---

5. From imagination to invention

The preceding model is imaginative but not inventive.

The noema Z was already part of the intentional language.

Genuine invention requires:

$$\boxed{Z \notin \qquad} \quad \mathcal{N}_C(t)$$

but later

$$\boxed{Z \in \mathcal{N}_C(t+1).}$$

The space of thinkable intentional objects itself changes.

Suppose C abstracts from its own temporal organization the operative law

$$L: \quad d' = m.$$

A generic variation operator can then produce:

$$\mathsf{Var}_{\neg}(L) = Z: \quad d' = \neg m.$$

The constructors existed before the invention; the resulting noema did not.

5.1 Why three intentional statuses suffice

For a finite inventor that can subsequently imagine and will the invented object, introduce:

$$u \in \{0, 1, 2\},$$

with:

$$u = 0: \quad Z \text{ unavailable},$$

$$u = 1: \quad Z \text{ invented and imagined},$$

$$u = 2: \quad Z \text{ willed}.$$

Using one trivalent variable is more economical than two independent bits for availability and attitude, because "willing a nonexistent noema" is not a well-typed state.

Define:

$$w(u) = \{0, \quad u = 0, 1, \ 1, \ u = 2.$$

and:

$$a' = a,$$

$$b' = a \oplus m \oplus w(u),$$

$$m' = a \oplus b.$$

Thus:

$$d' = m$$

until Z is willed, at which point:

$$d' = \neg m.$$

Under independence of the three statuses from the eight-state diachronic substrate:

$$8 \times 3 = 24.$$

Theorem 2. Minimal finite inventive-imaginative overmind

Under the assumptions that C:

1. begins without Z;
2. constructs Z endogenously;
3. can subsequently imagine Z;
4. can subsequently will that same Z;
5. allows the three statuses unavailable, imagined, and willed independently over its underlying diachronic configurations,

the state space has at least twenty-four states.

A twenty-four-state construction exists.

The important distinction is:

$$\boxed{\text{imagination changes the attitude toward an available possibility;}}$$

whereas:

$$\boxed{\text{invention changes the space of available possibilities.}}$$

---

6. Permanent invention and finite reversibility

The inventive model also exposes a structural incompatibility.

Suppose the repertoire variable changes:

$$Z \notin \mathcal{N} \longrightarrow Z \in \mathcal{N}$$

and the concept is never forgotten.

Then the acquisition is monotonic.

A closed finite reversible dynamical system is a permutation of a finite set. Every state belongs to a cycle.

A genuinely irreversible transition from “concept absent” to “concept permanently present” cannot lie on such a cycle.

Hence:

Lemma 3. Finite reversibility obstruction

A permanent monotonic concept acquisition cannot occur in a closed finite reversible dynamical system.

Consequently, a reversible physical realization must involve at least one of:

- environmental degrees of freedom carrying information about the pre-invention state;
- possible reverse trajectories corresponding to uninvention;
- an effectively unbounded state space.

This point becomes important for any future unitary quantum formulation.

---

7. Why no finite mind can invent forever

Suppose C can acquire:

$$Z_0, Z_1, Z_2, \ldots$$

with all Z_i remaining internally distinguishable.

If its complete relevant internal state space is finite, then by the pigeonhole principle some later stage must repeat an earlier state.

It can then no longer internally distinguish the richer later repertoire from the earlier one.

Thus:

Theorem 4. Finite-state generativity impossibility

No fixed finite-state monad can support indefinitely many internally distinguishishable novel noemata if their availability is encoded entirely in its internal state.

This forces the first qualitative jump in the project:

$$\boxed{\text{finite cognition} \longrightarrow \text{countably infinite generativity.}}$$

---

8. The minimal open-ended generative monad

The smallest infinite cardinal is:

$$\aleph_0.$$

A countably infinite system suffices.

Take one primitive noema:

$$Z_0$$

and one unary constructor:

$$S : \mathcal{N} \to \mathcal{N}.$$

Generate:

$$Z_{n+1} = S(Z_n).$$

Thus:

$$Z_0, \quad S(Z_0), \quad S^2(Z_0), \ldots$$

form the free unary term algebra:

$$\mathcal{T}_S(Z_0).$$

Its cardinality is exactly:

$$\aleph_0.$$

If S is non-collapsing,

$$S^n(Z_0) = S^m(Z_0) \Rightarrow n = m.$$

Hence an indefinitely growing family of distinct noemata is obtained.

For example:

$$Z_n = \text{“remain joint for } n \text{ intervals, then diverge.”}$$

The satisfaction conditions differ for different n, so this is not mere renaming.

Theorem 5. Minimal open-ended generative state cardinality

Under the assumptions that indefinitely many novel noemata must remain internally distinguishable and arise from a finite generative signature, the minimal state-space cardinality is:

$$\boxed{\aleph_0}.$$

It is attained by one primitive noema plus one non-collapsing unary constructor.

This suggests a minimal grammar:

$$\boxed{N ::= Z_0 \mid S(N).}$$

Call the resulting system GIM_1, the **Generative Intentional Monad of rank 1**.

---

9. Why cardinality ceases to measure cognitive complexity

Once:

$$|S| = \aleph_0,$$

adding finitely many noetic modes changes nothing cardinally:

$$2\aleph_0 = \aleph_0.$$

Likewise:

$$\aleph_0^2 = \aleph_0.$$

Thus a compulsory unary generator and a much richer deliberative system can have exactly the same cardinality.

From this point onward, an appropriate complexity signature should record something more like:

$$\mathfrak{C}(M) = (\text{control states}, \text{constructor signature}, \text{constructor arities}, \text{memory architecture}, \text{type depth}, \text{revision depth}$$

The project therefore crosses from a **state-counting regime** to a **structural complexity regime**.

---

10. Recombination requires branching structure

Unary generation produces an indefinitely long conceptual lineage but no independently variable branching.

For genuine recombination, require an output depending structurally on two independently chosen prior noemata.

A transparent free-term architecture with only nullary and unary operations cannot do this.

At least one binary constructor is needed.

Take one primitive atom z and one binary constructor:

$$B(-, -).$$

The grammar becomes:

$$\boxed{N ::= z \mid B(N, N).}$$

This is the free magma on one generator.

Its terms include:

$$z,$$

$$B(z, z),$$

$$B(B(z, z), z),$$

$$B(z, B(z, z)),$$

and indefinitely many branching structures.

Lemma 6. Binary recombination

Within a transparent free-term architecture, genuine recombination of two independently selectable noemata requires constructor arity at least two. One binary constructor suffices.

The state space remains countable.

---

11. Contrastive abstraction

Recombination alone is not abstraction.

We next require C to compare two intentional objects, identify an invariant structure across their difference, and retain the resulting invariant as a reusable operation.

Introduce a one-hole context:

$$\Box.$$

Schemas have grammar:

$$\boxed{K ::= \Box \mid B(K, N) \mid B(N, K).}$$

A schema such as:

$$\kappa = B(\Box, z)$$

acts on an ordinary noema:

$$\kappa[w] = B(w, z).$$

Suppose:

$$u = B(z, z)$$

and:

$$v = B(B(z, z), z).$$

They share the context:

$$B(\square, z).$$

Define a contrastive abstraction:

$$\boxed{\mathrm{Abs}(u, v) = B(\square, z).}$$

Conceptually, this is closely related to anti-unification: extracting a common structural generalization from two expressions.

11.1 Why two examples matter

From only one example, many abstractions are possible.

Given:

$$B(B(z, z), z),$$

one might extract:

$$\square,$$

$$B(\square, z),$$

or more specific contexts.

Variation between two presentations reveals what is invariant.

Thus, under a contrastive definition:

$$\boxed{\text{at least two co-accessible examples are required.}}$$

Husserlian retention supplies them economically:

$$R_C = u, \qquad I_C = v.$$

The already-existing temporal organization therefore becomes the memory architecture for abstraction.

---

12. Habituality as retained schema

The abstracted context should become reusable.

Let:

$$K_\kappa$$

denote the sedimented capacity associated with κ.

Then:

$$u, v \xrightarrow{\text{Abs}} \kappa$$

followed by:

$$\kappa, w \xrightarrow{\text{Plug}} \kappa[w].$$

This yields the intentional sequence:

$$\boxed{\text{experience} \to \text{contrast} \to \text{abstraction} \to \text{habitual schema} \to \text{application to new experience.}}$$

This is the first system in the hierarchy that does not merely create new intentional objects. It creates **new operations on intentional objects**.

Call it:

$$\boxed{\mathsf{ARM}_1}$$

for **Abstractive-Recombinational Monad**.

Its minimal transparent architecture contains:

- one primitive noema;
- one binary constructor;
- two co-accessible registers;
- contrastive abstraction;
- structural substitution.

All of these remain countable.

---

13. Open-ended schema learning

The learned schemas need not form a finite set.

Let:

$$r_0 = z,$$

$$r_{n+1} = B(r_n, r_n).$$

Define:

$$u_n = B(z, r_n),$$

$$v_n = B(B(z, z), r_n).$$

Then:

$$\mathrm{Abs}(u_n, v_n) = \kappa_n = B(\square, r_n).$$

Since:

$$r_n \neq r_m$$

for $n \neq m$,

$$\kappa_n \neq \kappa_m.$$

Hence C can learn indefinitely many reusable schemas:

$$\kappa_0, \kappa_1, \kappa_2, \ldots.$$

The noematic hierarchy has now become typed:

$$N_0 = \text{Noema},$$

$$N_1 \sim N_0 \to N_0.$$

The subject can intentionally manipulate **ways of thinking**, not only thought-objects.

---

14. Learning rules for constructing rules

The next transition introduces a second-order hole:

$$\Diamond.$$

A meta-schema takes a first-order schema as input.

For example:

$$\mu = B(\Diamond, z)$$

acts on:

$$\kappa = B(\Box, z)$$

to produce:

$$\mu[\kappa] = B(B(\Box, z), z).$$

Thus:

$$\boxed{\mu : (N \to N) \to (N \to N).}$$

Suppose C has independently learned:

$$\kappa_1 = B(\Box, z),$$

and:

$$\kappa_2 = B(B(\Box, z), z).$$

Comparison yields:

$$\mathrm{MetaAbs}(\kappa_1, \kappa_2) = B(\Diamond, z).$$

C can then extrapolate:

$$\kappa_3 = \mu[\kappa_2],$$

$$\kappa_4 = \mu[\kappa_3],$$

and so forth.

Before learning μ, each new κ_n might have required fresh object-level examples.

After learning μ, new rules can be generated from old rules.

This is genuine **meta-learning**.

---

15. A rank-2 intentional calculus

The hierarchy can be written:

Rank 0: noemata

$$N ::= z \mid B(N, N).$$

Rank 1: schemas

$$K ::= \square \mid B(K, N) \mid B(N, K).$$

Rank 2: meta-schemas

$$M ::= \Diamond \mid B(M, N) \mid B(N, M).$$

Operations include:

$$\mathrm{Abs} : N^2 \rightharpoonup K,$$

$$\mathrm{Plug} : K \times N \to N,$$

$$\mathrm{MetaAbs} : K^2 \rightharpoonup M,$$

and:

$$\mathrm{MetaPlug} : M \times K \to K.$$

Call this:

$$\boxed{\mathsf{MRM}_2},$$

the **Meta-Rule Monad of rank 2**.

The state space is still:

$$\aleph_0.$$

What has increased is type depth:

$$N_0,$$

$$N_1 = N_0 \to N_0,$$

$$N_2 = N_1 \to N_1.$$

This motivates a major general conclusion:

$$\boxed{\text{the meaningful complexity of advanced generative cognition lies in type structure rather than raw state-space car}}$$

---

16. Arbitrarily high finite meta-levels

Nothing forces the hierarchy to terminate at rank 2.

Define recursively:

$$N_{n+1} = N_n \to N_n.$$

Then rank $n+1$ contains transformations of rank-n operations.

For every finite n, the finite terms remain countable.

Even:

$$N_{<\omega} = \bigcup_{n<\omega} N_n$$

is countable.

Thus a monad can possess arbitrarily high finite orders of representation and metarepresentation while its formal state space remains:

$$\boxed{\aleph_0}.$$

This is another indication that cardinality has become almost irrelevant to cognitive power within the countable regime.

---

17. A deeper limitation: the abstraction mechanism remained fixed

Although MRM_2 can learn new rules and new rules for producing rules, one crucial process was still supplied externally:

$$\mathrm{Abs}\,.$$

The monad might learn infinitely many schemas while never changing its criterion of what counts as a meaningful structural similarity.

This motivates the next transition.

Let:

$$\Theta$$

denote a mutable **epistemic policy** specifying which structural transformations count as preserving relevant form.

Abstraction becomes:

$$\boxed{\mathrm{Abs}_\Theta\,.}$$

Changing:

$$\Theta_0 \to \Theta_1$$

therefore changes the abstraction mechanism itself.

---

18. The self-revising epistemic monad

Call the resulting architecture:

$$\boxed{\text{SREM}_1}$$

for **Self-Revising Epistemic Monad**.

Let Θ be represented as a finite rewrite or equivalence theory over the noematic term language.

Initially:

$$\Theta_0 = \varnothing.$$

Thus:

$$B(x, y) \not\equiv_{\Theta_0} B(y, x)$$

when $x \neq y$.

Suppose, however, C repeatedly encounters paired structures:

$$p = B(x, y),$$

$$q = B(y, x)$$

whose relevant outcomes are indistinguishable:

$$\omega(p) = \omega(q).$$

The current epistemic policy distinguishes them structurally while experience supplies evidence of a systematic invariance.

C constructs the hypothesis:

$$r_{\text{swap}} : B(X, Y) \equiv B(Y, X).$$

Then:

$$\Theta_1 = \Theta_0 \cup \{r_{\text{swap}}\}.$$

Under the new policy:

$$B(u,v) \equiv_{\Theta_1} B(v,u)$$

for arbitrary u, v.

Thus:

$$\boxed{\mathrm{Abs}_{\Theta_0} \ /} \qquad = \mathrm{Abs}_{\Theta_1} .$$

C has changed what counts as a pattern.

---

19. Evidence-driven policy revision

A candidate epistemic revision should not be adopted merely because it can explain one coincidence.

Introduce a score:

$$L(\Theta; D) = \mathrm{Err}(\Theta; D) + \lambda\,\mathrm{Complexity}(\Theta),$$

where D is accumulated evidence.

Then:

$$\Theta_{t+1} = \begin{cases}\Theta^*, & L(\Theta^*; D_t) < L(\Theta_t; D_t),\ \Theta_t,\ \text{otherwise.}\end{cases}$$

The exact score is merely one toy implementation.

The structural criterion for genuine self-revision is:

$$\boxed{\exists u, v : \mathrm{Abs}_{\Theta_{t+1}}(u,v)\ /} \qquad = \mathrm{Abs}_{\Theta_t}(u,v).$$

If no future abstraction changes, the revision is epistemically cosmetic.

The system can also retract or refine a rule if later evidence shows that the proposed invariance was too broad.

Hence development can take the form:

$$\boxed{\text{generalization} \to \text{exception} \to \text{differentiation} \to \text{revised ontology.}}$$

---

20. Retrospective reinterpretation

A revised Θ can be applied not only to future experience but also to retention.

An earlier retained presentation:

$$R$$

originally synthesized according to Θ_0 can be reconsidered under:

$$\Theta_1.$$

Thus:

$$\mathrm{Reinterpret}_{\Theta_1}(R)$$

may transform:

> “a different structure”

into:

> “another presentation of the same invariant structure.”

The temporal phenomenology therefore becomes:

$$\boxed{\text{new epistemology} \to \text{new past-as-understood} + \text{new future horizon.}}$$

This gives retention a third role in the research program.

It first supported subject persistence, then supplied contrasts for abstraction, and now supplies material for **revision of one's own former understanding**.

---

21. Epistemic autobiography

Add the self-noema:

$$\sigma_C.$$

The monad can represent:

$$\mathrm{UsesPolicy}(\sigma_C, \Theta_0),$$

and later:

$$\mathrm{UsesPolicy}(\sigma_C, \Theta_1).$$

It can therefore form:

$$\mathrm{ChangedPolicy}(\sigma_C, \Theta_0, \Theta_1),$$

and perhaps:

$$\mathrm{ChangedBecause}(\sigma_C, \Theta_0, \Theta_1, D).$$

This yields the intentional content:

> I used to organize these presentations one way; evidence led me to organize them differently.

This is a minimal form of **epistemic autobiography**.

---

22. But the evaluator was still fixed

The self-revising epistemic monad can alter:

$$\Theta,$$

but not the criterion:

$$L$$

according to which alternative policies are compared.

Thus it has changed its inductive bias while retaining a fixed meta-standard of epistemic success.

The next step places even that standard inside the revisable constitution.

Let:

$$\boxed{\mathcal{E} = (\Theta, \ell)}$$

denote the epistemic constitution, where:

$$\Theta$$

specifies structural/inductive policy and:

$$\ell$$

specifies the current epistemic evaluator.

Now permit:

$$\boxed{(\Theta_t, \ell_t) \to (\Theta_{t+1}, \ell_{t+1}).}$$

The monad can therefore revise what it counts as a better epistemology.

---

23. The threat of infinite evaluative regress

If:

$$\ell_0$$

evaluates the proposed transition to:

$$\ell_1,$$

one might ask what justifies ℓ_0.

Introducing:

$$\ell_2$$

merely generates:

$$\ell_0, \ell_1, \ell_2, \ldots$$

without grounding the sequence.

The alternative developed here is a **reflective closure condition** rather than an infinite tower.

A predecessor and successor epistemology jointly constrain the transition.

---

24. Trivial self-endorsement is insufficient

Any constitution $\mathcal{E}$ could contain an evaluator that declares:

$$\mathcal{E}$$

uniquely optimal.

Therefore:

Lemma 7. Trivial self-endorsement

Unconstrained self-endorsement cannot provide nontrivial epistemic justification.

A candidate successor cannot be licensed solely because its newly installed evaluator approves of itself.

Some continuity with the predecessor and with evidence must remain.

---

25. Bilateral reflective ratification

Suppose:

$$\mathcal{E}_0 = (\Theta_0, \ell_0)$$

and candidate:

$$\mathcal{E}_1 = (\Theta_1, \ell_1).$$

A simple initial model requires:

$$\ell_0(\mathcal{E}_1; D) < \ell_0(\mathcal{E}_0; D)$$

and:

$$\ell_1(\mathcal{E}_1; D) < \ell_1(\mathcal{E}_0; D).$$

Thus the predecessor says:

> I have grounds for leaving my present constitution.

The successor says:

> given the transition history, I endorse arriving here.

This is **diachronic reflective equilibrium** rather than hierarchical evaluation.

A fixed point is an epistemic constitution for which no admissible candidate transition is mutually preferred.

---

26. Why global cross-evaluation is still too strong

Radical conceptual change may produce epistemologies whose full vocabularies are mutually incommensurable.

It may be meaningless for:

$$\ell_0$$

to evaluate every detail of $\mathcal{E}_1$.

A weaker and more flexible notion is therefore needed.

Instead of demanding complete mutual evaluation, require a **local justificatory bridge**.

---

27. The bridge witness

Let:

$$w$$

be an internally representable witness recording:

$$(\mathcal{E}_0, D, \mathcal{E}_1, \mathcal{R}),$$

where $\mathcal{R}$ is the relevant transition rationale or provenance.

A valid epistemic transition requires w to be ratifiable from both sides:

$$\mathrm{Rat}_{\mathcal{E}_0}(D, w),$$

and:

$$\mathrm{Rat}_{\mathcal{E}_1}(D, w).$$

Thus:

$$\boxed{\mathcal{E}_0 \xrightarrow[D,w]{} \mathcal{E}_1.}$$

The epistemologies need not share a global language.

They need only possess enough overlapping structure for the passage to be intelligible as justified from predecessor and successor perspectives.

This transforms epistemic continuity from global commensurability into **local bridgeability**.

---

28. The reflectively self-grounding epistemic monad

Call the resulting system:

$$\boxed{\mathrm{RSGEM}_1}$$

for **Reflectively Self-Grounding Epistemic Monad**.

Its intentional state may be written:

$$C_t = (E_C, \sigma_C, R_t, I_t, P_t, \mathcal{E}_t, D_t).$$

Its epistemic constitution can include much more than (Θ, ℓ):

$$\boxed{\mathcal{E} = (\mathcal{O}, \Lambda, \Theta, \ell, \mathcal{V}, \ldots),}$$

where:

- $\mathcal{O}$ is ontology;
- Λ is a consequence relation;
- Θ is structural/inductive policy;
- ℓ is epistemic evaluation;
- $\mathcal{V}$ contains simplicity, predictive, explanatory, or other epistemic values.

All of these may change.

Hence even:

$$\boxed{\text{logic itself may be historically revisable.}}$$

The remaining question is what, if anything, must stay fixed.

---

29. Minimizing the invariant kernel

An earlier version of the model retained several substantive constraints: consistency, evidence responsiveness, nontriviality, computability, and so forth.

Closer examination suggests that much of this can be moved into the revisable constitution.

Classical consistency is unnecessary because a monad could use a paraconsistent logic.

A fixed notion of truth is unnecessary.

Bayesian probability is unnecessary.

A simplicity principle is unnecessary.

Even a fixed scalar loss function is unnecessary.

This leaves a much thinner candidate kernel.

---

30. K1: reflective addressability

For genuinely reflective self-revision, C must be able to make its own epistemic organization an intentional object.

There must be something like:

$$\boxed{\text{Quote} : \mathcal{E} \mapsto \ulcorner \mathcal{E} \urcorner.}$$

Without this, an epistemic policy may adapt, but the subject cannot intentionally represent:

> my method of forming beliefs is changing.

Thus:

$$\boxed{\mathrm{K1}: \ \text{Reflective addressability}.}$$

The constitution need not possess unrestricted semantic self-reference. It only needs internally usable access to descriptions of itself and its candidates.

---

31. K2: nondegenerate coupling to givenness

The process must also remain genuinely epistemic rather than arbitrary self-modification.

Let D represent what is given:

- perception;
- memory;
- failed prediction;
- proof;
- contradiction;
- surprise;
- action outcome;
- internal anomaly.

The kernel does not dictate what counts as evidence in detail.

It requires only that givenness can make a difference.

Schematically, there must exist cases for which:

$$\mathcal{E}_1 \prec_{D_1} \mathcal{E}_2$$

while:

$$\mathcal{E}_1 \not\prec_{D_2} \mathcal{E}_2.$$

Thus epistemic transition is not wholly determined by self-certification.

Call this:

$$\boxed{\text{K2 : Nondegenerate evidence coupling.}}$$

Phenomenologically it is a minimal principle of **receptivity**.

---

32. K3: diachronic bridgeability

Finally, successor epistemologies must not legitimate themselves solely retrospectively.

A transition:

$$\mathcal{E}_0 \to \mathcal{E}_1$$

requires an addressable witness:

$$w$$

that has standing from both sides.

Define:

$$\boxed{\mathsf{Bridge}(\mathcal{E}_0, D, \mathcal{E}_1, w).}$$

This entails sufficient provenance for the same continuing subject to retain:

- what its former epistemology was;
- what relevant givenness occurred;
- what epistemology succeeded it;
- what justified the passage.

Thus:

$$\boxed{\text{K3 : Diachronic bridgeability.}}$$

This is the epistemic analogue of the identity-transport principles that appeared near the beginning of the monadology.

---

33. The candidate minimal kernel

The resulting proposal is:

$$\boxed{K_{\min} = \{K1, K2, K3\}.}$$

Or in phenomenological language:

$$\boxed{K_{\min} = \{\text{self-reference}, \text{receptivity}, \text{retention}\}.}$$

The claim is not that these constitute a finished theory of rationality.

It is that they may form the thinnest invariant architecture within which radical epistemic self-revision still remains recognizable as **reflectively epistemic**.

---

34. Independence-style arguments

The three conditions appear individually necessary within the present definition.

34.1 Remove K1

Without reflective addressability:

$$\mathcal{E}_0 \to \mathcal{E}_1$$

may still be learning or adaptation.

But the system cannot take its own epistemology as an intentional object.

Hence:

$$\boxed{\neg K1 \Rightarrow \text{adaptation without reflective self-revision.}}$$

34.2 Remove K2

If transition is wholly independent of givenness, epistemic change becomes internal mutation, schedule, or self-certification.

If everything is ratified equally, ratification has no discriminatory content.

Hence:

$$\boxed{\neg K2 \Rightarrow \text{self-change without epistemic responsiveness.}}$$

34.3 Remove K3

Without a predecessor-successor bridge, a candidate can install an evaluator that simply declares itself correct.

There is no diachronic justificatory continuity.

Hence:

$$\boxed{\neg K3 \Rightarrow \text{potentially arbitrary criterion drift.}}$$

These are not yet universal impossibility proofs. They establish relative independence with respect to the phenomena as presently defined.

---

35. The kernel as a structural diagram rather than a doctrine

An important conceptual shift follows.

The invariant kernel need not be a set of propositions accepted by every epistemology.

It may instead be the **shape of admissible transformation**.

A revision has the structure:

$$\boxed{D \qquad \swarrow \qquad \searrow \quad \mathcal{E}_0 \quad \xleftrightarrow{w} \quad \mathcal{E}_1 \quad \searrow \qquad \swarrow \qquad \sigma_C}$$

The subject remains the same diachronic intentional continuant while its substantive epistemology changes.

This resembles a span:

$$\mathcal{E}_0 \longleftarrow W \longrightarrow \mathcal{E}_1,$$

where W is a space of locally shared justificatory witnesses.

This suggests a future categorical treatment in which epistemologies are objects and admissible justificatory bridges are morphisms, spans, or higher cells.

---

36. Radical conceptual change and the epistemic Ship of Theseus

Consider:

$$\mathcal{E}_0 \xrightarrow{w_0} \mathcal{E}_1 \xrightarrow{w_1} \mathcal{E}_2 \xrightarrow{w_2} \cdots .$$

After many revisions:

$$\mathcal{E}_n$$

may share little substantive content with:

$$\mathcal{E}_0.$$

Its ontology, logic, standards of simplicity, and theory of evidence may all have changed.

Yet each transition can remain locally bridgeable.

Define historical epistemic continuity:

$$\mathcal{E}_0 \sim_{\text{hist}} \mathcal{E}_n$$

when a finite chain of admissible bridges connects them.

This produces an **epistemic Ship of Theseus**.

Identity of the knowing subject need not require invariance of its worldview.

Instead:

$$\boxed{\text{radical conceptual transformation may preserve rational continuity through locally justified diachronic transport.}}$$

---

37. The return of Husserlian temporality

At first glance the later stages of the program may seem far removed from the original retention–impression–protention structure.

In fact the same structure repeatedly reappears.

At the eight-state level:

$$R - I - P$$

constitutes the first genuinely diachronic overmind.

For abstraction:

$$R + I$$

provide contrastive examples.

For epistemic revision:

$$R$$

preserves the former interpretation against which the present anomaly becomes visible.

For reflective self-grounding:

$$R - I - P$$

can become:

$$\boxed{\text{former epistemology} - \text{present epistemology/anomaly} - \text{possible revised epistemology.}}$$

Thus temporal synthesis is not merely one early phenomenological decoration.

It becomes the architecture supporting increasingly sophisticated forms of cognition.

This suggests:

$$\boxed{\text{learning itself may be intrinsically diachronic in a stronger sense than mere storage of previous data.}}$$

The present is intelligible as a possible revision of the retained past toward a protended future organization.

---

38. The emerging hierarchy

The entire development can now be summarized according to **what the monad is capable of changing**:

weak collective	:	current relational state,
diachronic overmind	:	state inherited through collective history,
self-conscious overmind	:	intentional relation to its own stream,
imaginative overmind	:	attitude toward unrealized possibilities,
inventive overmind	:	set of available intentional objects,
generative monad	:	unbounded sequence of intentional objects,
abstractive monad	:	repertoire of operations on intentional objects,
meta-rule monad	:	operations for constructing operations,
self-revising epistemic monad	:	criteria by which patterns are discovered,
reflectively self-grounding monad	:	criteria by which epistemic criteria are evaluated.

The deepening structure is therefore:

$$\boxed{\text{state} \to \text{object} \to \text{possibility} \to \text{new object} \to \text{generator} \to \text{schema} \to \text{meta-schema} \to \text{inductive bias} \to \text{epistemi}}$$

---

39. Complexity transitions

The mathematical character of minimality changes along this hierarchy.

For the earliest stages, state cardinality distinguishes architectures:

$$4, \quad 8, \quad 16, \quad 24.$$

Open-ended invention forces:

$$\aleph_0.$$

After that, all systems considered remain countable.

Their relevant complexity is instead measured by:

Constructor arity

Unary versus binary generation.

Type rank

$$N_0,\ N_1,\ N_2, \ldots.$$

Memory/workspace structure

How many co-accessible intentional objects or schemas are required?

Generative closure

Can the system construct indefinitely many objects, rules, or meta-rules?

Revision depth

Which parts of its own cognitive architecture remain fixed?

This motivates replacing one scalar “complexity of mind” by a structured invariant:

$$\boxed{\mathfrak{C}(C) = (c, a, r, m, g, e),}$$

where, schematically:

- c: control complexity;
- a: maximal constructor arity;
- r: representational type rank;
- m: memory/workspace architecture;
- g: generative depth;
- e: epistemic revision depth.

Developing this invariant systematically is an important next step.

---

40. Relation to overmind subjecthood

The later generative systems were described abstractly as “monads,” but they can be grafted back onto the original interacting-overmind architecture.

Take the eight-state substrate:

$$(a, b, m).$$

Attach a countably generative intentional register:

$$N,$$

and later a mutable epistemic constitution:

$$\mathcal{E}.$$

A full state may therefore look like:

$$\boxed{(a, b, m, N, \mathcal{E}, q).}$$

Because a finite product with a countable set is countable:

$$|\{0,1\}^3 \times \mathbb{N} \times \cdots| = \aleph_0.$$

So the original overmind can, in principle, be enriched all the way through the hierarchy without changing its state-space cardinality beyond countability.

The constituents A and B can remain distinct subjects throughout.

The result is a nested collective subject that can:

- retain collective history;
- recognize itself;
- imagine;
- invent;
- abstract;
- construct meta-rules;
- revise its own inductive bias;
- revise its epistemic standards;
- represent the history of those revisions.

Thus the finite overmind construction was not an unrelated curiosity. It supplies a plausible **minimal temporal subject-kernel** onto which the later cognitive architecture can be built.

---

41. What these results do not establish

Several cautions remain essential.

None of these formal models proves that a corresponding automaton is phenomenally conscious in our world.

Subjecthood remains supplied by a psychophysical realization law:

$$\mathcal{P}_W.$$

The mathematical constructions establish conditional structural statements of the form:

> If a possible psychophysical world associates subjecthood with processes satisfying the stipulated criteria, then the following architecture is sufficient or minimally sufficient under the stated independence conditions.

Similarly, the finite-state numbers are not universal consciousness constants.

The eight-, sixteen-, and twenty-four-state results depend on explicit modeling assumptions.

The claims concerning unary and binary grammars are minimal within a **transparent typed term architecture**. Opaque encodings, universal combinators, Gödel coding, or unconventional representational schemes can compress several functions into one primitive.

The significance is therefore diagnostic:

$$\boxed{\text{the models expose conceptual dependencies that richer implementations can obscure.}}$$

---

42. New conjectures suggested by the development

Several broader conjectures have emerged.

42.1 Temporal-Closure Conjecture

Independent diachronic state is a deeper transition toward subjecthood than explicit self-reference.

The recurring importance of retention supports the hypothesis that mature subjecthood is fundamentally processual rather than instantaneous.

42.2 Generative-Mind Conjecture

Advanced intentional agency is characterized not merely by motion within a fixed state space but by endogenous expansion or restructuring of its own intentional possibilities.

42.3 Type-Depth Conjecture

Once open-ended generativity is reached, representational type depth and transformability are more informative measures of cognitive complexity than cardinality.

42.4 Epistemic Plasticity Conjecture

A sufficiently reflective subject need not possess permanently fixed substantive principles of ontology, inference, induction, or epistemic evaluation. What must remain invariant may be much thinner.

42.5 Minimal Epistemicity Conjecture

For a reflectively self-transforming intentional subject, the irreducible conditions of epistemic agency may be only:

$$\boxed{\text{reflective self-addressability} + \text{receptivity to givenness} + \text{diachronic justificatory continuity.}}$$

All recognizably substantive epistemological principles may be revisable realizations built above this kernel.

---

43. Research program

The present sequence now suggests a much more concrete next phase.

43.1 Prove independence of K_1, K_2, K_3

Construct three pathological monads:

$$K_2 + K_3 + \neg K_1,$$

$$K_1 + K_3 + \neg K_2,$$

and:

$$K_1 + K_2 + \neg K_3.$$

Determine formally which capacities survive.

This would turn the current minimal-kernel argument into an actual independence analysis.

43.2 Test whether one kernel axiom is derivable

Perhaps suitable forms of reflective addressability and bridgeability already imply some weak evidence coupling.

Or perhaps bridgeability can be reconstructed from self-addressability plus temporal retention.

Finding such dependencies could shrink:

$$K_{\min}$$

still further.

43.3 Formalize epistemic bridges categorically

Define a category, bicategory, or double category whose:

- objects are epistemic constitutions;
- 1-morphisms are justified revisions;
- 2-morphisms compare bridge witnesses.

Composition should answer whether:

$$\mathcal{E}_0 \to \mathcal{E}_1 \to \mathcal{E}_2$$

induces a coherent bridge:

$$\mathcal{E}_0 \to \mathcal{E}_2.$$

This would connect epistemic persistence to the earlier coarse-graining and identity-transport ideas.

43.4 Develop the complexity invariant $\mathfrak{C}$

Classify intentional systems not merely by state-space size but by:

$$(\text{arity}, \text{type depth}, \text{memory}, \text{generativity}, \text{revision depth}).$$

A partial order of cognitive architectures may emerge.

43.5 Formalize abstraction through anti-unification

Our use of one-hole contexts should be connected rigorously to anti-unification and related term-rewriting machinery.

This could yield actual algorithms and minimality theorems rather than hand-selected examples.

43.6 Study self-modifying abstraction policies computationally

Enumerate finite rewrite policies Θ over small term languages and determine:

- which evidence streams trigger revision;
- whether revision cycles;
- whether stable equilibria exist;
- whether concept formation becomes more efficient;
- whether pathological self-certification arises.

43.7 Return to reversible and quantum realizations

The eight- and sixteen-state systems admitted reversible implementations.

Permanent invention introduced an irreversibility obstruction.

A quantum version should ask whether generative and epistemically revising monads can be represented as open quantum systems with memory while the larger closed system evolves unitarily.

43.8 Investigate collective epistemic revision

If C is an overmind composed of A and B, its epistemic constitution may disagree with those of its members.

Then we can ask:

$$\Theta_A, \Theta_B, \Theta_C$$

and:

$$\ell_A, \ell_B, \ell_C.$$

Can collective evidence cause C to revise while A and B do not?

Can member dissent act as epistemic evidence for the overmind?

Can the higher subject teach a revised epistemology downward?

This would reconnect the later epistemic theory directly to the societies-of-mind program.

43.9 Study epistemic fission and fusion

What happens when a subject with one epistemic history divides into two successors?

Can:

$$\mathcal{E}_C$$

produce:

$$\mathcal{E}_A \ / \ \mathcal{E}_B =$$

while both inherit the same justificatory past?

Conversely, how can two epistemic histories fuse without simply discarding their conflicts?

The bridge formalism may supply an especially natural language for these questions.

43.10 Search for the true bedrock

The most immediate target remains:

$$\boxed{\text{Is } K_{\min} = \{K_1, K_2, K_3\} \text{ genuinely minimal?}}$$

The answer could be yes.

But perhaps reflective epistemicity has an even smaller structural essence.

---

44. Conclusion

The developments reported here began with an apparently modest question: how many states are required for the smallest intentional overmind?

That question led to a sequence of structural thresholds.

An eight-state system can sustain a diachronic collective state irreducible to an instantaneous description of its members. Self-indexing need not enlarge that state space. An independent reflective or imaginative attitude introduces an additional binary distinction, yielding sixteen states. Genuine finite invention introduces a third intentional status and a natural twenty-four-state model.

At that point finite-state methods reach a boundary.

A mind that can continue inventing indefinitely must possess an unbounded representational capacity. Countably infinite term systems suffice. The smallest such system requires only a primitive intentional object and a non-collapsing constructor.

Richer cognitive powers then arise not by increasing cardinality but by changing the architecture of representation.

A binary constructor permits recombination.

Contrastive comparison produces abstractions.

Retained abstractions become habitual schemas.

Schemas themselves become objects of comparison, producing meta-rules.

The abstraction mechanism becomes parameterized by a mutable epistemic policy.

The epistemic policy becomes self-revisable.

Finally, the standards used to evaluate policy revision themselves become part of the revisable epistemic constitution.

At this point the threat of infinite regress appears. The proposed resolution is not an infinitely ascending hierarchy of evaluators but a diachronic network of locally shared justification. Predecessor and successor epistemologies are linked by retained bridge witnesses constrained by givenness and belonging to one continuing subject.

This in turn suggests that remarkably little need remain invariant.

Perhaps not classical logic.

Perhaps not a fixed ontology.

Perhaps not Bayesianism, simplicity, or one permanent theory of truth.

Perhaps not even a fixed standard of epistemic improvement.

The tentative residue is:

$$\boxed{\text{self-addressability} + \text{receptivity} + \text{diachronic bridgeability.}}$$

This would make the deepest epistemic structure of the monad strikingly similar to the deepest phenomenological structure discovered much earlier in the program.

The eight-state overmind became subjectlike when the collective present acquired a history of its own.

The self-grounding epistemic monad becomes rationally continuous when its present worldview remains intelligibly related to the worldview from which it arose.

Both results point toward the same general principle:

$$\boxed{\textbf{the identity, intentionality, creativity, and rationality of a mature monad may be fundamentally pr}}$$

The original problem of interacting Husserlian monads has therefore expanded into a broader mathematical question:

$$\boxed{\text{What are the minimal structures by which a subject can change not only its states and beliefs, but its concepts, its}}$$

The models developed here provide a first hierarchy of answers.

References

[1] *Possible Worlds of Interacting Husserlian Monads: Toward a Panprotopsychist Process Algebra of Subject Formation, Fusion, Fission, Overminds, and Collective Imagination.* Unpublished research manuscript developed in the present project.

[2] *Minimal Finite Models of Diachronic and Self-Conscious Overminds: Eight-State and Sixteen-State Theorems in a Toy Husserlian Monadology.* Unpublished technical note developed in the present project.