A Structural Model of Depression Recovery as Constrained Aviplestos

A Technical Blueprint

Huỳnh Mai Phúc (Plone Mraz)

Independent Researcher

Vietnam, 2026

ORCID: 0009-0009-0571-7151

TECHNICAL BLUEPRINT — WORKING DRAFT *Not peer-reviewed. Circulated for critique and collaboration. Contains no clinical recommendations.*

Epistemic Status (read this before anything else)

This document occupies a deliberately limited epistemic position, and that position is stated up front so that no reader has to infer it.

This is a speculative theoretical hypothesis. It is not supported by data collected under this framework. No parameter in the model below has been estimated. No prediction in this document has been tested. The model has not been fitted to any clinical dataset, and the author has not conducted the empirical work that would be required to evaluate it. A reader who treats any statement here as an established result has misread the document.

What is offered is narrower and, the author hopes, more honest: a **structural framework** that (i) reinterprets a recognised clinical phenomenon — recovery in depression under combined pharmacological and cognitive intervention — through the lens of a prior philosophical framework (*Aviplestos*; Huỳnh Mai Phúc, 2026b), and (ii) derives from that reinterpretation **at least one falsifiable prediction that the standard models do not by themselves generate**. The value of the framework, if it has any, will be decided by whether that prediction survives empirical test — not by the internal elegance of the model.

A scope restriction is also stated up front. An earlier version of this corollary extended the same reasoning to spontaneous remission in oncology and to congenital genetic disease. **That extension has been removed in full and is not defended anywhere in this document.** The reasons are given in Section 7; in brief, that extension failed on both epistemological grounds (selection on the dependent variable) and ethical grounds (it reproduced the structure of a clinical narrative that has been empirically tested, largely

unsupported, and shown to impose harm on patients who do not recover). The present document is restricted to depressive disorder, where a genuine empirical anchor exists.

Contribution and Scope (what this is, what it is not, who it is for)

This section states, before the argument begins, exactly what this document contributes and what it does not. It is load-bearing, not prefatory: a reader who skips it will mis-rank everything that follows. The contribution divides into two parts with deliberately different epistemic standing, plus an explicit boundary of authorship. The first part must not be allowed to borrow the credibility of the second, and neither may borrow the credibility of work not yet done.

What this is not. This document contributes no new treatment, no change to clinical practice, no dosing or therapeutic instruction, and no claim that any patient outcome would differ if a clinician read it. It contains no data, no estimated parameters, and no analysed case. Any reading of it as "a new philosophy-derived treatment for depression" is a misreading the document explicitly rejects, and such a reading would be correctly dismissed.

Contribution A — conditional empirical contribution (value zero until tested). There is exactly one potential empirical contribution, and it is confined to Section 5: the prediction that, once the pharmacological barrier is lowered, the *directional alignment* of an intervention with the patient's gap predicts response more strongly than the *intensity or dose* of that intervention. The standard accounts — pharmacological dose–response and Behavioral Activation — do not, combined, by themselves generate this direction-dominates-magnitude asymmetry for the post-pharmacological window. That asymmetry is, in principle, measurable. Its honest present value is stated without softening: **on current evidence this contribution has zero realised value.** It has only conditional, potential value, and the condition is the completion of Gate 1 (closing the citation gap) and Gate 2 (specifying the falsification experiment) in the validation roadmap, neither of which has been done. Contribution A is a promissory note, and it is labelled as one.

Contribution B — framework contribution (independent of data). Three contributions do not depend on any dataset and are realisable now. First, the document converts a recognised but informally-handled clinical risk — the early-phase return of drive before the return of life-directed orientation — into a quantity internal to a model: the Risk Window Theorem gives that known risk a structural address (the lag between the two equations' time constants) rather than treating it as an experiential caution. Second, it supplies a shared vocabulary across pharmacological and psychotherapeutic intervention, which currently describe their interaction in separate languages — explicitly a *terminological* contribution, not a demonstrated mechanism, and ranked as such. Third, it imposes a falsifiability discipline on a domain prone to unfalsifiable metaphor, by forcing

the intuition that "a single sentence can produce a disproportionate recovery" into a statement that can be wrong. These are real contributions at the rank of *conceptual framework*, not *discovery*; the document does not permit B to be read at the rank of A.

The boundary of single-author completion. This document is the work of an independent researcher without clinical specialisation, institutional affiliation, or funding, and the division of labour this implies is stated rather than hidden. The work that one person in that position can complete is Gate 1 and Gate 2: the literature review and the design of the falsification experiment — conceptual and synthetic work. The work that one person in that position *cannot* complete is not a matter of effort. Gate 4 (independent ethical review, with a safety protocol for the predicted self-harm window) is, by construction, not grantable to an unaffiliated individual conducting research on an at-risk population without clinical supervision — and it should not be; the Risk Window Theorem is precisely what makes that gate non-circumventable. Gate 5 (pilot data and parameter estimation) requires four distinct professional competencies — clinical psychiatry, psychometric measurement, biomedical statistics, research governance — plus participant infrastructure and funding. No single person is those four roles, regardless of depth. This is a structural boundary, not a deficiency of the author.

The consequence is not that the contribution is void; it is that the contribution has a *shape*. What an independent researcher without institutional means can produce is a falsifiable framework, stated precisely enough that others who do hold clinical access, ethical oversight, and funding can take it up and test it. That is the function this document is built to serve, and it is the reason the document's stated purpose is transfer to others for collaboration and critique. The boundary between what is complete here and what requires an institution is not an apology for the work; it is part of the work, and stating it plainly is a condition of the document's honesty.

Table of Contents

Epistemic Status (read this before anything else) 1
Contribution and Scope (what this is, what it is not, who it is for) 2
1. The Clinical Risk Window (stated first, by design) 5
2. Position Within the Research Programme 6
3. The Core Reinterpretation 7
3.1 The Patient Has Not Lost Aviplestos 7
3.2 The Role of Medication: Barrier-Lowering, Not Motivation-Injection 7
3.3 The Role of Ignition: A Cognitive Event at a Lowered Barrier 7
4. The Formal Model 9
4.1 The Unified Vector Equation (formal statement) 9
4.2 Symbol Glossary (plain-language, every term) 9
4.3 The Decomposition Corollary (where the risk window comes from) 11
4.4 Inferring the Gap from an Emission-Channel Observation Space (with explicit noise model) 12
5. The Discriminative Prediction 15
5A. Test Design for the Discriminative Prediction (Gate 2) 16
The author's boundary, stated plainly 17
6. Convergence: Behavioral Activation and the Activation Phenomenon 19
7. What Was Removed, and Why 20
8. Computational Realisation Path (prospective, untested) 21
9. Human-Subjects and Data Ethics Preconditions (blocking; open) 22
10. Validation Roadmap (not a product roadmap) 24
11. Limitations 25
References 26

1. The Clinical Risk Window (stated first, by design)

This section is placed before the model rather than after it. The placement is intentional. Any framework that describes the *re-energising* of a depressed patient without, in the same breath, describing the *danger* of re-energising a patient is structurally incomplete and ethically negligent. The danger is not a caveat appended to the model. On the present account it is a **theorem of the model** — it falls out of the mathematics in Section 4.3 and cannot be separated from it.

The phenomenon is independently recognised in clinical psychiatry and is not original to this framework. Recovery from depression is not uniform across dimensions. Psychomotor energy, initiative, and the capacity to act frequently return *before* mood, cognition, and the reorganisation of forward-directed goals return. This is sometimes discussed in connection with the early phase of antidepressant response. The clinically significant consequence is that there can exist a window in which a patient has regained the capacity to act but has not yet regained a life-directed orientation for that capacity.

The Aviplestos framework does not soften this. It sharpens it. *Aviplestos*, in its founding statement (Huỳnh Mai Phúc, 2026b), is explicitly **value-neutral**: it is the structural tendency of an entity toward configurations it does not yet occupy, with no built-in commitment to those configurations being good. A value-neutral vector, re-activated inside a cognitive system whose model of its own available configurations is distorted, will move toward whatever configuration that distorted system reads as the open one — and for a depressed cognition, that configuration can be self-termination rather than re-engagement.

Stated in the framework's own terms: **the danger is not that the patient's Aviplestos is absent. The danger is precisely that it is present, value-neutral, and re-activated before its direction has reconverged on a life-directed configuration.** Section 4.3 shows that this window is not an empirical add-on but a direct consequence of decomposing the model into magnitude and direction.

No clinical instruction is given anywhere in this document. The risk window is presented as a structural feature of the model, addressed to researchers and reviewers, not as guidance to any practitioner or patient.

2. Position Within the Research Programme

This blueprint is a corollary branch of *Aviplestos: The Structural Tendency Toward the Not-Yet* (Huỳnh Mai Phúc, 2026b). It does not restate that paper's argument; it assumes it. The single load-bearing import is the founding definition: Aviplestos is the structural tendency of any entity toward configurations it does not yet occupy — a vector with magnitude and direction — observable not through direct perception of "the gap" but through the entity's actions of filling it, and carrying no normative content.

The corollary asks a bounded question: *if* Aviplestos is taken as a valid descriptive lens, what structural account does it give of the clinically observed pattern in which pharmacological and cognitive interventions, combined, sometimes produce recoveries from depression that are disproportionate to either intervention alone? The answer offered is a model in which the depressed state is not the loss of the vector but the locking of it behind a barrier, and recovery is the lowering of that barrier followed by the directional re-convergence of the vector.

The corollary is **not** claimed to be more fundamental than, or a replacement for, any clinical or psychological account of depression. Its relationship to the existing evidence base is examined in Section 6.

3. The Core Reinterpretation

3.1 The Patient Has Not Lost Aviplestos

The reinterpretation rests on a single move. A depressed patient is not modelled as an entity from which the vector toward life, filling, and connection has been *removed*. The vector is modelled as *present but locked*: held in place by an activation barrier — neurochemical and cognitive — that is large enough that the system cannot, unaided, traverse it.

(An analogy used in earlier drafts — a diamond locked in its crystal lattice, freed by heat — has been discarded. It is physically self-defeating: heating diamond drives it toward graphite, a *lower*-order configuration, which inverts the intended meaning. The framework uses the language of reaction kinetics directly instead: an activation barrier lowered by a catalyst-like intervention that is not itself consumed and does not itself supply the forward tendency. This is not a decorative change; it is the difference between a metaphor that contradicts the model and a structure that the model can actually be written in.)

3.2 The Role of Medication: Barrier-Lowering, Not Motivation-Injection

On the standard folk picture, treatment "adds" motivation to the patient from outside. The present model rejects this picture for the pharmacological component. Antidepressant medication, in this model, does **not** add to the forward tendency. It **lowers the activation barrier** so that the tendency the patient already has can traverse it.

This is why the barrier term must enter the model multiplicatively, inside an exponential gate, and not as a subtractive term. A subtractive barrier produces an unphysical prediction: as the barrier grows without bound, the rate of recovery would fall to negative infinity — unbounded decay. The phenomenology of locked depression is not unbounded decay; it is *freezing* — the vector held near zero throughput. An Arrhenius-type gate produces exactly this: as the barrier grows without bound, the gate closes toward zero and the vector's throughput freezes, rather than reversing into infinite decay. The model is written this way because it is the only form in which the sentence "the drug lowers the barrier, it does not inject the drive" is *true of the equation* and not merely asserted beside it.

3.3 The Role of Ignition: A Cognitive Event at a Lowered Barrier

Once the barrier is lowered, the model holds that the vector still requires a directional input — an "ignition." A single utterance, a psychotherapeutic intervention, a perspective-changing event: each is modelled not as a source of energy added to the patient but as a directional input that orients an already-present, now-traversable vector. This is the precise sense in which "a single sentence" can be associated with a disproportionate recovery: not because the sentence is energetically large, but because, at a lowered barrier, the

determining variable is *alignment with the gap*, not *magnitude of input*. Section 5 turns this into a falsifiable claim.

4. The Formal Model

The model is presented at full mathematical detail, and every symbol is then explained in plain language with no mathematical prerequisite. Readers who do not work with differential equations can read Section 4.2 alone and lose none of the conceptual content.

4.1 The Unified Vector Equation (formal statement)

Aviplestos for the patient is a vector, written $\vec{A}$. A vector is a single object carrying two pieces of information at once: how large it is (magnitude) and where it points (direction). The formal statement of the model is a single equation governing this single object:

$$\frac{d\vec{A}}{dt} = \underbrace{k\, e^{-E_a/T}}_{\text{activation gate}} \cdot \underbrace{\rho(\vec{I},\vec{G})}_{\text{resonance}} \cdot \underbrace{f(\tau)}_{\text{temporal profile}} \cdot \underbrace{\hat{G}}_{\text{gap direction}} - \underbrace{\gamma\, \vec{N}_{\text{env}}}_{\text{environmental loss}} - \underbrace{\nabla V(\vec{A})}_{\text{bistable + saturation structure}}$$

This single equation is the formal statement of the corollary. It is deliberately presented as one equation, because in the founding paper Aviplestos *is* one vector, not two separate quantities; presenting it as one object is the form that is faithful to the philosophy. The decomposition into two equations in Section 4.3 is not a competing model — it is this same equation, rewritten so that a clinically critical feature, otherwise hidden inside the single object, becomes visible.

4.2 Symbol Glossary (plain-language, every term)

Every symbol above, with no assumed mathematical background:

- $\vec{A}$ **— the patient's Aviplestos.** A single arrow. Its *length* is how much life-directed drive is currently in throughput; its *heading* is what configuration that drive is currently pointed at (re-engagement, or self-termination, or somewhere between).
- $\frac{d\vec{A}}{dt}$ **— the rate of change of that arrow over time.** Note carefully: this is the *velocity* of recovery (how fast the arrow is changing right now), not the *acceleration*. Earlier drafts labelled it "acceleration"; that was an error of naming and has been corrected. Nothing in the model is a second derivative.
- E_a **— the activation barrier.** How hard it currently is for the vector to move at all. High in locked depression. Lowered — not removed — by pharmacological intervention.
- T **— the system's effective flexibility** (a "temperature," by analogy with reaction kinetics). How much capacity the neurocognitive system has to traverse the barrier. Pharmacological intervention can be modelled equivalently as lowering E_a *or* raising T

; mathematically these are the same effect, which is itself a structural claim worth testing.

- $e^{-E_a/T}$ **— the activation gate.** The single most important structural choice in the model. When the barrier is very high relative to flexibility, this number is near zero and the whole forward term is switched off — the vector *freezes*. When the barrier is lowered, the gate opens and the forward term can flow. This is the mathematical form of "the drug unlocks; it does not push."
- $\vec{I}$ **— the interaction input.** Whatever is delivered into the patient's space: a therapeutic utterance, an intervention, a relationship, an event. It has a direction (what it is pointed at) and a size.
- $\vec{G}$ **— the gap.** The configuration the patient does not yet occupy but structurally could — in the relational sense defined in the founding paper, not a stored target. $\hat{G}$ is its direction alone.
- $\rho(\vec{I}, \vec{G})$ **— the resonance function.** This measures *how well the input is aligned with the gap*, independent of how large the input is. It is defined as the cosine of the angle between the input and the gap: it is $+1$ when the input points exactly along the gap, 0 when it is perpendicular (irrelevant), and negative when the input points against the gap (counter-therapeutic). **Note on notation:** earlier drafts wrote this as θ. That was a poor choice, because θ conventionally denotes an *angle*, not the cosine of an angle. It is renamed ρ here. This correction is adopted from external critique and is acknowledged as correct.
- $f(\tau)$ **— the temporal profile of the input.** Whether the input is delivered as a single burst or sustained over time. Reserved here as a named function; its specific form (impulse vs. sustained vs. decaying) is left open as a modelling parameter, not silently fixed.
- $\vec{N}_{\text{env}}$ **— environmental load.** The patient's environment, which can act as noise (draining) or, with opposite sign, as support (assisting). γ converts it into the same units as the rest of the equation so the subtraction is well-defined.
- $V(\vec{A})$ **— the potential landscape**, and $\nabla V(\vec{A})$ **— its slope.** This is the term that does two jobs. First, it encodes **saturation**: there is a configuration of full recovery beyond which the vector does not keep growing without bound (a boundary condition earlier drafts omitted). Second, and critically, it encodes a **bistable structure**: the landscape has two low points — one "life-directed" basin and one "self-terminating" basin. Which basin the vector's *heading* falls into is the entire content of Section 1, now located precisely inside the mathematics.

A note on honesty about this object: $V(\vec{A})$ is, at this stage, a **structural placeholder, not a measured function**. Asserting that the landscape is bistable is a hypothesis, not a finding.

The model states *that* the landscape must have this shape for the framework to be coherent with the known clinical risk window; it does not claim to have established the shape empirically. This is flagged here rather than buried.

4.3 The Decomposition Corollary (where the risk window comes from)

The single vector $\vec{A}$ can always be written as a length multiplied by a heading: $\vec{A} = A\,\hat{u}$, where A is "how much drive" and $\hat{u}$ is "pointed where." Rewriting the single equation in these two parts does not change the model; it makes one feature legible that the single form hides.

Magnitude (how much drive there is):

$$\frac{dA}{dt} \;=\; k\, e^{-E_a/T}\, \rho(\vec{I},\vec{G})\, f(\tau)\, \underbrace{A\left(1 - \frac{A}{A_{\max}}\right)}_{\text{saturation}} \;-\; \gamma\, N_{\text{env}}$$

Direction (where that drive is pointed):

$$\frac{d\hat{u}}{dt} \;\propto\; \rho(\vec{I},\vec{G})\,\left(\hat{G}_{\text{life}} - \hat{u}\right) \;-\; \nabla_{\hat{u}} V(\hat{u})$$

The first equation says the *amount* of drive can be rising. The second equation says the *heading* of that drive moves toward the life-directed configuration only as fast as the resonance term ρ can turn it, against the pull of the bistable landscape.

The clinically critical consequence now follows as a structural fact, not an empirical caveat:

> **Risk Window Theorem (informal).** There exists an interval during which $\frac{dA}{dt} > 0$ — drive is being restored — while $\hat{u}$ has not yet converged on $\hat{G}_{\text{life}}$ — the heading has not yet reconverged on a life-directed configuration. During this interval the patient's restored drive is, by the model's own structure, capable of being directed at the self-terminating basin of V.

This is the same window described in Section 1 and independently recognised in clinical psychiatry. The point of the decomposition is that the framework **predicts the window from its own structure** rather than having to apologise for it afterward. A model in which re-energising could not, in principle, precede re-orientation would be a model that contradicts a known clinical fact. This one does not, and it locates exactly where the danger lives: in the gap between the two equations' time constants.

4.4 Inferring the Gap from an Emission-Channel Observation Space (with explicit noise model)

Every equation so far contains the term $\vec{G}$, the gap. Sections 4.1–4.3 use it freely. This section confronts the question those sections deferred: **if the gap cannot be observed directly, how is the $\vec{G}$ that appears in every equation actually obtained?** The construct introduced here is native to this blueprint. It is derived from the internal commitments of the Aviplestos framework and is not imported from, nor dependent on, any other framework; the motivation below stands entirely on Aviplestos's own definitions.

The construct follows from two prior commitments, not from a new assumption. First, the founding framework defines the gap as a *differential* — the difference between an entity's current configuration and the configuration space its governing constraints make available (Huỳnh Mai Phúc, 2026b,). Second, the same framework establishes that this differential is not directly observable, even in principle; it is a structural feature of the observation relation itself, and can only be inferred from what the entity *emits* (ibid., §3.1–3.2). The black-hole image used there is exact: the mass is never seen; it is reconstructed from the radiation of the matter around it. The question this section answers is the one that image forces — *what is the accretion disk, here?*

Definition (emission-channel observation space). A depressed entity does not emit "its gap." It emits, across several distinguishable channels, observable proxies of where it is currently held short of the configurations otherwise available to it. Define the **deficit-signal vector**

$$\vec{D} = \left(D_{\text{som}},\ D_{\text{cog}},\ D_{\text{mean}},\ D_{\text{eff}},\ D_{\text{rel}}\right)$$

whose components are the five channels along which a depressed entity emits an observable deficit-proxy. Each is defined here natively, by its clinical emission, not by analogy to anything external:

- D_{som} **— somatic channel.** Emitted bodily depletion: sleep architecture disruption, psychomotor change, energy and appetite signals. The most directly measurable channel.
- D_{cog} **— cognitive channel.** Emitted processing distortion: ruminative narrowing, attentional capture, degraded signal-to-noise in self-relevant inference.
- D_{mean} **— meaning channel.** Emitted collapse of significance: anhedonia of purpose, loss of forward-directed value, the flattening of what the entity treats as worth moving toward. This channel carries most of the directional information about $\hat{G}$, and least of its magnitude.
- D_{eff} **— effort channel.** Emitted capacity to expend effort toward recovery: the observable cost the entity can currently bear to traverse the activation barrier. **A required value-neutrality clause attaches to this channel and is not optional.** D_{eff}

is a structural measure of *available effort-capacity under load*, not a measure of the patient's character, willingness, or moral effort. A low D_{eff} is a reading of system state — it is, in depressive disorder, a core symptom (avolition), not a deficiency of will the patient is answerable for. Any clinical deployment of this model inherits this clause explicitly; the value-neutrality of the variable does not survive translation into a clinical register on its own and must be carried, in words, wherever D_{eff} is used.

- D_{rel} **— relational channel.** Emitted decoupling: withdrawal, breakdown of internal and social synchrony, the loss of the entity's coupling to the systems it ordinarily co-regulates with.

The channels are presented as five because that is the resolution at which a depressed entity's emission is clinically distinguishable, not because five is structurally necessary. The model does not claim this basis is complete or orthogonal; channel correlation and missing channels are absorbed by the noise model below, not denied.

The inference map and its irreducible noise. The gap is then estimated, never measured, by an inverse map from the emitted deficit-signal to the latent differential:

$$\vec{G} \approx \Phi(\vec{D}) + \varepsilon, \qquad \varepsilon = \varepsilon_{\text{model}} + \varepsilon_{\text{omit}}\left(\sum_c 1[\text{channel } c \text{ unobserved}]\right)$$

Here Φ is the (non-invertible, lossy) reconstruction from emitted channels to the latent gap; $\varepsilon_{\text{model}}$ is the irreducible approximation error the founding framework already commits to (the observer's model of the gap is always an approximation, never a replica; ibid., §3.2); and $\varepsilon_{\text{omit}}$ is an explicit omission term that *grows with each channel left unobserved*. The second term is the formal discharge of a commitment made informally elsewhere in this programme — that any incompleteness in data collection propagates into the gap estimate. It is written here so that the incompleteness is a named quantity in the model, not a verbal caveat outside it.

Three honesty constraints on this subsection, stated rather than buried. First, Φ **is a structural placeholder, not an estimated function.** Its existence is asserted; its form is not known and has not been fitted. Naming Φ does not constitute solving it. Second, **the channel basis is a modelling choice, not a discovered partition.** A different clinically-motivated basis is possible; the claim is only that *some* finite emission basis plus an explicit omission term is the honest structure, not that this particular five is privileged. Third, **this section does not make the gap observable.** It makes explicit the inference by which an unobservable gap is approximated from observable emission, with the error of that inference carried as a term rather than disowned. The gap remains, as the founding framework requires, something only ever estimated under irreducible uncertainty.

This subsection is the operational precondition for what follows. Section 5's discriminative prediction concerns alignment between an intervention and the gap; that prediction is only testable once the gap has a defined — if noisy — inference procedure. That procedure is

the content of this subsection, and its noise model is the reason the prediction in Section 5 must be stated as a statistical asymmetry rather than a deterministic law.

5. The Discriminative Prediction

The recurring objection to every branch of this research programme is the non-discrimination problem: a framework that is consistent with every outcome explains none. This corollary is required to earn its place by generating at least one prediction that the standard models do not by themselves generate. It does.

The resonance function ρ is defined as alignment *independent of input magnitude* (the cosine, not the projection). This is a deliberate modelling commitment, and it has a sharp consequence:

> **Direction-dominates-magnitude (testable).** At a sufficiently lowered barrier (post-medication, gate open), the model predicts that the *directional alignment* of an intervention with the patient's gap will predict response **more strongly than the intensity or dose of the intervention**. A small, well-aligned input ($\rho \approx 1$, small $|\vec{I}|$) should outperform a large, poorly-aligned input ($\rho \approx 0$, large $|\vec{I}|$).

A pure dose–response model and a pure behavioral-activation account do not, on their own, predict that *alignment* dominates *intensity* once the pharmacological gate is open. This framework does. That asymmetry is, in principle, measurable, and it is the proposed point of empirical contact for the entire corollary.

A competing variant — replacing the cosine with the projection $\text{Proj}_{\vec{G}}(\vec{I})$, which reintroduces input magnitude — is *not* dismissed. It is more biologically realistic (a vanishingly small input, however well-aligned, will not clear the barrier). But adopting it silently would erase the discriminative prediction and collapse the model back toward a generic dose–response account. The two variants are therefore presented as a **head-to-head empirical question**, not a modelling preference: the cosine model is retained as the falsifiable core; the projection model is named as the realistic alternative; and which one fits is itself the experiment, not a choice to be made at the desk.

5A. Test Design for the Discriminative Prediction (Gate 2)

This section specifies how the prediction in Section 5 could be falsified. It is the completion of Gate 2 in the validation roadmap. It is written to one standard: precise enough that a researcher with clinical and statistical authority can act on it or reject it on sight. It does not attempt to be a protocol; it is the logic a protocol would be built on, plus an explicit statement of where the author's competence ends.

The core design problem. The prediction refers to two quantities that cannot be measured directly: the gap $\vec{G}$ (Section 4.4 establishes it is only ever inferred, with irreducible noise) and the directional alignment $\rho(\vec{I}, \vec{G})$ that depends on it. A naive design that measures ρ and correlates it with response would fail — not because the prediction is wrong, but because ρ cannot be measured cleanly. The design must therefore avoid measuring ρ in absolute terms and instead manipulate it in relative terms only.

The core idea, in one sentence. Do not measure alignment; manipulate its *ordering*, set that ordering deliberately *against* intensity, and rest the entire test on a single inequality.

The logic has four pillars.

Pillar 1 — the pharmacological gate is an inclusion criterion, not a variable. The prediction is stated only for the post-gate-open state. Only patients who have already reached a minimum pharmacological-response threshold are admitted. Below that threshold the prediction does not apply, and mixing such patients in destroys the test by noise.

Pillar 2 — individualise direction, then sever direction from intensity. For each patient, the dominant deficit channel is identified from the emission-channel space of Section 4.4. "Aligned" means an intervention targeting that dominant channel; "misaligned" means an intervention targeting a non-dominant but still clinically legitimate channel. The test never needs absolute ρ — only the within-patient ordering "more aligned / less aligned," which is far more controllable.

Pillar 3 — the 2×2 intensity-inversion grid. This is the heart of the discrimination:

	Low intensity	**High intensity**
Aligned	cell A — the decisive cell	cell B
Misaligned	cell C	cell D

A pure dose–response account predicts D > B > C > A: intensity dominates, the high cell always beats the low cell. The Aviplestos prediction asserts something different and incompatible: **A > D** — the aligned-but-low-intensity cell outperforms the misaligned-but-high-intensity cell. Cell A is exactly where the two models contradict each other. An

observed A > D, after controlling for the pharmacological gate, is an observation the dose–response model cannot produce. The whole test reduces to that one inequality.

Pillar 4 — intensity must be operationalised orthogonally to direction. If "intensity" inadvertently correlates with "alignment," the 2×2 collapses. Intensity should be a direction-neutral dosage — for example duration, frequency, or number of sessions of the same intervention type — such that intensity can be raised without making the intervention more aligned. Keeping these two axes orthogonal is the single strictest design requirement.

A second discrimination, almost free. Section 5 leaves open whether ρ is pure cosine (direction only) or projection (direction × magnitude). The same 2×2 answers this without a separate study: if pure cosine holds, intensity adds little within the aligned row, so cell A ≈ cell B; if projection holds, intensity still adds even when aligned, so cell B > cell A markedly. The A-versus-B comparison is the cosine/projection discriminator, reusing the data already collected for the main test.

Three ways this design can be wrong, stated before they are asked. First, a *circularity risk*: the emission-channel space of Section 4.4 is used to define the dominant channel, which defines "aligned," which is then used to test a prediction about alignment. If the channel inference is wrong, "aligned" is wrong, and the test tests the wrong thing. This is mitigated, not eliminated, only by determining the dominant channel with an instrument independent of the framework being tested (e.g. blinded clinical assessment), so that "direction" is not defined by the very framework under test. Second, an *ethical obstruction*: deliberately delivering a known-misaligned intervention to one arm, in a population with self-harm risk (the Risk Window), is not a neutral design choice — it is a serious research-ethics question, and it belongs to Gate 4, not Gate 2. This design *assumes* something Gate 4 may forbid; that is flagged here, not hidden inside the diagram. Third, the entire test hangs on the *response measure* — which scale, measured when, by whom (blinded). That is a psychometric question and a load-bearing one.

The author's boundary, stated plainly

This subsection is written without mitigation, and nothing after it walks it back.

The design above was produced by an author without clinical specialisation, and proposed in dialogue with an AI system that also has no clinical-trial design expertise. Two parts of the design have different standing, and they are not blended.

What the author can stand behind is the *logical architecture*: the reduction of the prediction to the single inequality A > D; the within-patient control; the inversion of intensity against direction; the cosine-versus-projection discrimination falling out of the A-versus-B comparison. This part is assessed by reasoning consistency, not by clinical credential, and it is offered for that kind of scrutiny.

What the author cannot stand behind, and does not: every clinical parameter. Sample size. Which validated instruments. The pharmacological-response threshold defining gate-open. How to operationalise psychotherapeutic "intensity" cleanly. The independent instrument for dominant-channel identification. The response measure and its timing. These are not gaps the author can close with more effort; they require an authority the author does not have, and supplying invented values for them would be the precise failure this document was built to refuse. They are left as named absences, not filled with plausible-looking numbers.

There is no claim attached to this boundary. It is not "limited, but valuable nonetheless." It is the statement of where the author's competence ends, full stop. The design is offered to those who hold the missing authority as something to act on or reject — not as something complete. Whether it has worth is for that reader to decide, and stating that plainly is not modesty; it is the only honest description of what this section is.

6. Convergence: Behavioral Activation and the Activation Phenomenon

The depression branch is defended, and the oncology/genetic branch was not, for one reason: this branch has a real empirical anchor, and that branch had only analogy.

Behavioral Activation is an established, evidence-based treatment for depression whose core principle is that structured action precedes and produces improvement in mood, rather than improved mood being a precondition for action. This is, in clinical-trial form rather than philosophical form, the same structural claim as the "epistemological reversal" in the founding paper (Huỳnh Mai Phúc, 2026b): action is prior; the gap becomes legible through the acting. This is presented as a **convergence**, in the precise sense the founding paper uses that word: not a source the framework was derived from, but an independently established body of evidence that describes what appears to be the same structure. The convergence is what distinguishes this branch from the removed one.

A second convergence: the early-phase return of psychomotor energy and initiative ahead of mood and cognition is independently recognised in psychiatry. The present model predicts exactly this ordering from its structure — the activation gate opens on the magnitude equation before the direction equation has reconverged — which is also precisely the mechanism that generates the Risk Window Theorem of Section 4.3.

Two clarifications are required for honesty. First, convergence with Behavioral Activation supports the claim that *the structure is real and clinically salient*. It does **not** by itself support the claim that *this particular formalisation is the correct description* of it. Those are different claims with different burdens, and only the discriminative prediction in Section 5 bears on the second. Second, this section does not claim that the Aviplestos formalism is more fundamental than Behavioral Activation, nor that Behavioral Activation is a special case of it. They are two descriptions, in different languages, of what may be the same structure.

7. What Was Removed, and Why

This section is retained in the document deliberately, as a record of a deletion, because the deletion is itself part of the intellectual content.

An earlier version extended this corollary to spontaneous remission in cancer and to survival in congenital genetic disease, under the heading of a "biological Aviplestos" tending toward self-healing. That extension has been removed in full, on two independent grounds, either of which is sufficient.

Epistemological. The extension selected on the dependent variable. It examined cases that *did* recover and labelled the recovery as Aviplestos, while having no falsifiable account of the cases that did not. A framework under which both recovery and non-recovery confirm the framework predicts nothing. This is the non-discrimination problem in its most acute and least defensible form.

Ethical. The claim that a will-like or tendency-like inner orientation determines survival through serious organic disease has been empirically examined in oncology and has largely failed to find that psychological orientation determines prognosis; the "fighting spirit" narrative has been specifically criticised for imposing a cruel secondary burden — the implication that patients who do not survive failed to want it enough. Renaming "will" as "Aviplestos" does not change this structure; it relabels a narrative that has been tested, largely unsupported, and shown to cause harm. A separate sentence in the earlier draft — that congenital genetic disease might be survived "without depending much on medical treatment" — is one inferential step from a refusal-of-care rationale, and a framework is judged by how it can be read and used, not only by its author's intent.

The depression branch is retained because it does not have either defect: it has a falsifiable prediction (Section 5) and a real empirical anchor (Section 6). The author intends to revisit the biological question only after acquiring the relevant medical literature and an account of non-recovery cases that is falsifiable rather than post hoc. Until then, it is not defended here.

8. Computational Realisation Path (prospective, untested)

This section describes how the model **would be realised computationally in order to be tested further** — not how a system or product would be built. There is no system. There is no product. The only computational object this framework currently warrants is an instrument for subjecting the model in Section 4 to falsification. Every statement in this section is conditional on the discriminative prediction of Section 5 surviving at least one empirical test; absent that, none of the work below is justified.

The realisation path, conditional on that precondition, has three components.

First, **numerical integration of the coupled system.** The decomposition in Section 4.3 is a pair of coupled differential equations — one for magnitude A, one for heading $\hat{u}$ — with a bistable potential V. These cannot be solved in closed form and would require numerical integration to expose their behaviour, in particular the *duration and parameter-dependence of the Risk Window* identified in Section 4.3. The first concrete computational deliverable is therefore not an application; it is a simulation that makes the risk window quantitatively visible as a function of the model's parameters, so that the window can be reasoned about before any human data exists.

Second, **parameter estimation against data, once data exists.** The parameters k, E_a, T, γ, the form of $f(\tau)$, and the shape of V are currently unmeasured (Section 11). A realisation path would specify how each maps to an observable quantity and how it would be estimated. This component is explicitly **blocked** by Section 9: no estimation against human data may occur until the preconditions in that section are satisfied, and at the present time they are not.

Third, **a discriminative-test harness.** Section 5 frames cosine-resonance versus projection as a head-to-head empirical question. The computational realisation includes a procedure that, given a dataset, would discriminate between the two variants rather than fit one and assume it. This is the component that protects the framework from the non-discrimination problem at the level of implementation, not only at the level of theory.

No technology stack, framework, language, or infrastructure is specified here, deliberately. Specifying one would imply a build that is not warranted at this epistemic stage and would misrepresent a speculative model as an engineering project. The realisation path is a description of *what computational work would falsify the model*, ordered so that the risk-window simulation comes first and any human-data component comes last and gated.

9. Human-Subjects and Data Ethics Preconditions (blocking; open)

Decision status, stated plainly. The author has *not* decided whether this model will ever be taken to the collection of real-world depression data from human participants. This section is therefore written as a set of **blocking preconditions**, not as a compliance plan. Its function is to make explicit what would have to be true *before* any such step, so that the question cannot be crossed silently. Nothing in this document authorises, schedules, or assumes the collection of human data. If that line is ever approached, this section is the gate, and at the present time the gate is closed.

This is deliberately *not* framed as "data security and compliance standards." That framing would treat the central issue as technical (encryption, access control), when the central issue is **research ethics involving human subjects in a population at elevated risk**. Data security is real and is included below, but it is a subordinate clause, not the main one.

The blocking preconditions, any one of which being unmet is sufficient to halt:

First, **independent ethical review.** No data from human participants may be collected without prior approval from a competent institutional review board or research ethics committee. An independent researcher without institutional affiliation does not thereby acquire an exemption from this requirement; the absence of an institution is a reason the requirement is *harder* to satisfy, not a reason it does not apply.

Second, **informed consent appropriate to the population.** Consent procedures must be designed for participants experiencing depressive disorder, including consideration of how depressive cognition affects comprehension and voluntariness. Generic consent boilerplate is not sufficient for this population.

Third — and this is the precondition that is specific to *this* model and follows directly from its own structure — **a safety protocol for the Risk Window.** Section 1 and Section 4.3 establish, from the model's own mathematics, that there can exist an interval in which a participant's drive is restored before its direction has reconverged on a life-directed configuration. A framework that predicts this window from its structure cannot ethically be taken near human participants without a concrete, pre-registered safety protocol addressing exactly that window: monitoring during the predicted interval, escalation pathways, and the involvement of qualified clinical personnel. The model is not permitted to describe the danger on paper and then approach human data without an operational answer to it. This is the single precondition the author cannot delegate, defer, or treat as a formality.

Fourth, **data minimisation and protection, as a subordinate technical layer.** Within and only within the boundary established by the three preconditions above: collect the minimum data the test requires; store it encrypted; restrict access; define retention and

deletion in advance. This is necessary but it is the *last* item here, not the first, because solving it does not make the prior three any less binding.

The honest summary: this document does not meet any of these preconditions, and is not currently on a path that requires meeting them. They are recorded here so that the transition from "speculative model" to "research on people" cannot happen without first passing, visibly, through this section.

10. Validation Roadmap (not a product roadmap)

This is a roadmap of *evidence*, not of *features*. It contains no versions, no releases, no milestones expressed as deliverables. It is a sequence of **gates**, each of which is a blocking condition: until a gate is cleared, the gates after it do not exist. The ordering is itself a claim — that the cheap, safe, non-human work must exhaust itself before the expensive, risk-bearing, human-facing work is even considered.

Gate 1 — Close the citation gap. The literature review deferred throughout this document (see the Note on References) is completed: the clinical claims in Sections 1, 6, and 7 are either substantiated against the published literature or withdrawn. This gate is first because it is the cheapest, carries no risk, and may by itself reveal that the framework's empirical anchor is weaker or stronger than assumed. No subsequent gate is justified until this one is cleared.

Gate 2 — Specify the discriminative experiment. The direction-dominates-magnitude prediction of Section 5 is turned from a statement into a concrete experimental design: what is manipulated, what is measured, what result would confirm and — critically — what result would falsify it, including the cosine-versus-projection discrimination. This is a paper-and-design gate; it involves no participants.

Gate 3 — Risk-window simulation. The computational realisation of Section 8, component one, is built: the coupled system is integrated numerically and the Risk Window of Section 4.3 is characterised quantitatively as a function of parameters. The window is understood *in silico* before it is ever approached *in vivo*.

Gate 4 — Ethics preconditions. Every blocking precondition in Section 9 is satisfied, in full, with the Risk-Window safety protocol foremost. This gate is where the author's *undecided* status (the answer that generated this section) is resolved one way or the other — and it cannot be resolved by default or by drift. It is resolved explicitly or the roadmap stops here permanently.

Gate 5 — Pilot data and parameter estimation. Only after Gates 1–4 are cleared: a pilot dataset is collected under the approved protocol, parameters are estimated, and the Section 5 prediction is tested for the first time against real data.

The roadmap is deliberately front-loaded with the gates that cost nothing and risk no one, and the only gate that touches a human being is last and is preceded by an explicit, non-defaultable decision. If the roadmap halts at Gate 1 or Gate 2 — which is a real possible outcome — the framework remains exactly what the top of this document says it is: a speculative theoretical hypothesis, honestly labelled, that did not survive contact with the evidence. That outcome is not a failure of the roadmap; it is the roadmap working.

11. Limitations

Beyond the global epistemic status stated at the top, the following specific limitations are acknowledged.

First, **no parameter has been estimated and no prediction has been tested.** k, E_a, T, γ, the form of $f(\tau)$, and above all the shape of the potential V are unmeasured. The bistability of V is assumed, not demonstrated.

Second, **the formalisation is a structural sketch, not a validated dynamical system.** It has not been checked for the existence, uniqueness, or stability of its solutions, and no claim is made that the chosen functional forms are the correct ones rather than the simplest ones consistent with the conceptual content.

Third, **the non-discrimination problem is contained, not eliminated.** Section 5 provides one discriminative prediction. One prediction is the minimum, not a sufficiency. Until it is tested, the framework's claim to explanatory rather than merely re-descriptive status is unestablished — and is stated here as unestablished, not glossed.

Fourth, **the convergence argument cuts both ways.** That independent observers (clinical and philosophical) converge on an action-prior structure supports the reality of the phenomenon, not the correctness of this description of it. The document tries to respect that distinction throughout; readers should hold it to that standard.

Fifth, **this is not clinical guidance and must not be read as any.** It contains no protocol, no dosing logic, no therapeutic instruction, and no recommendation regarding any treatment decision. It is a theoretical document addressed to researchers and reviewers. The Risk Window Theorem is included specifically so that the model cannot be read as describing re-energisation without describing its danger; it is not, and must not be used as, clinical advice.

References

- Huỳnh Mai Phúc. (2026b). *Aviplestos: The Structural Tendency Toward the Not-Yet*. Working Paper. https://doi.org/10.57967/hf/11444
- Huỳnh Mai Phúc. (2026a). *The Synthetic Principle: An Observation of the Mechanism by Which Discrete Elements Generate Irreducible Configurations*. Working Paper. https://doi.org/10.57967/hf/11444

Note on references. This blueprint cites the author's own programme papers for the imported definitions. Clinical claims referred to in Sections 1, 6, and 7 (the early-phase activation pattern in antidepressant response; Behavioral Activation as an evidence-based treatment; the empirical literature on psychological orientation and cancer prognosis) are stated at the level of well-established clinical knowledge and are deliberately *not* given specific citations here, because the author has not yet conducted the literature review that responsible citation would require. Supplying citations not yet verified would violate the standard this document holds itself to. Those citations are an explicit open task before any version of this document is circulated beyond a critique context.