THEORETICAL FRAMEWORK PAPER · 2025

The Upstream Signal Model: *A Predictive Processing Framework* for Emotional Experience, Symptomatic Presentation, and the Problem of Downstream Intervention

An updated theoretical account including original contributions on constitutional constraint architecture, mechanistic emotion classification, identity-layer rigidity, and the problem of flexibility as a clinical target

Sarah Martinez, LCSW
Inside Attunement
insideattunement@gmail.com · upstreammap.netlify.app

ABSTRACT

The Upstream Signal Model (USM) is a theoretical framework for understanding emotional experience, nervous system activation, and symptomatic presentation as outputs of a predictive, resource-managing system — rather than as discrete states to be identified, categorized, and managed. The model synthesizes established findings from predictive processing theory (Clark, 2016; Friston, 2010), constructed emotion theory (Barrett, 2017), allostatic regulation (Sterling, 2012), somatic experiencing (Levine, 1997), and complexity science (Deacon, 2012) into an integrated mechanistic architecture.

The central claim is that what we call an emotion — and what diagnostic nosology calls a symptom — is not a primary event but a downstream compression artifact: the most interpreted, most filtered output of a cascade that begins at the level of prediction error and moves through body signal, arousal, strategy, feedback, concept, and narrative before arriving at the word.

This updated version of the paper extends the original USM architecture with four theoretical additions: (1) a four-domain constitutional constraint model — Viability, Bonds, Autonomy, Orientation — that specifies the domains the nervous system is organized to protect prior to conscious awareness; (2) a proposed four-type mechanistic emotion classification as an illustration of what mechanistic granularity could look like in practice; (3) a structural account of identity-layer rigidity and why it is systematically misidentified as flexibility in clinical and spiritual frameworks; and (4) an account of organic drift toward prediction-defending in adult systems and its implications for the design of interventions.

SECTION 1

Introduction

The predominant models of emotional life used in clinical practice — DSM diagnostic categories, cognitive-behavioral formulations, affect regulation frameworks — share a structural assumption: that emotions are discrete, identifiable states that can be detected, named, and addressed at the level at which they present. A person is anxious, depressed, shameful, or dysregulated. The intervention targets that state.

This assumption is not wrong, exactly. Emotions as experienced are real. The problem is epistemological: when we name the output of a complex generative process and treat that name as the primary unit of clinical analysis, we build interventions on a highly compressed representation of whatever is actually driving the presentation. We are working at the mouth of the river while the source runs unchecked upstream.

The Upstream Signal Model proposes an alternative architecture — one grounded in predictive processing (Clark, 2016; Friston, 2010), emotion construction theory (Barrett, 2017), and anticipatory regulation (Sterling, 2012). In this architecture, emotional experience is not detected but assembled; symptoms are not discrete pathologies but mechanistically distinct system states that produce similar surface presentations; and intervention that targets the label without reaching the generating mechanism will be, at best, partially effective.

This paper presents the theoretical foundations of the USM, describes its original structural contributions, and identifies the specific ways in which it extends, reframes, or departs from its source literature. The updated version incorporates theoretical extensions developed through clinical application and theoretical development following the original formulation.

SECTION 2

Theoretical Foundations

2.1 Predictive Processing and Active Inference

The predictive processing framework — developed primarily by Clark (2013, 2016) and Friston (2010) — proposes that the brain is fundamentally a prediction machine. Rather than passively receiving sensory input, the brain continuously generates top-down predictions about the causes of sensory signals and uses incoming data primarily to compute and minimize prediction error: the discrepancy between what was anticipated and what arrived.

Friston's extension into active inference further proposes that action is itself a form of prediction error resolution: the organism can minimize prediction error either by updating its model (inference) or by acting on the world to make it match the prediction (action). The USM adopts this bidirectionality as foundational, and extends it to include a third route: acting on the self-model to absorb prediction error, which looks like updating but is mechanically closer to action. This distinction is developed in Section 5.

2.2 Constructed Emotion Theory

Barrett's theory of constructed emotion (Barrett, 2017) challenges the classical view of emotions as hardwired, universal responses. In Barrett's account, emotions are constructed in real time from interoceptive signals, prior conceptual knowledge, and contextual meaning. The same physiological state can become anxiety or excitement depending on the concepts available and the context in which they are deployed.

Barrett's emotional granularity finding further shows that individuals who possess finer-grained emotional concepts show measurably different physiological regulation, health outcomes, and capacity for effective action. The USM adopts this framework and extends it upstream: the construction process is treated as clinically secondary to what initiates it. The USM additionally proposes a reinterpretation of the granularity finding — that the regulatory benefit may derive not from the precision of the label itself but from the act of pre-conceptual inquiry the labeling task requires. This is addressed in Section 5.2.

2.3 Allostasis and the Body Budget

Sterling (2012) introduced allostasis as a revision to homeostatic models. Where homeostasis proposes a fixed setpoint, allostasis proposes that the body sets physiological parameters anticipatorily — based on predicted future demands. Allostatic load is not the cost of normal adaptation but the cost of chronic, unresolved, or miscalibrated mobilization. The USM adopts this framing and translates it into the tank construct — a clinical metaphor for allostatic reserve as the finite resource available at any moment.

2.4 Somatic Experiencing

Levine's somatic experiencing framework (Levine, 1997, 2010) proposes that trauma is stored in the body as incomplete motor sequences and survival responses that were interrupted before reaching biological resolution. The USM adopts the clinical primacy of pre-linguistic body signal and the incomplete-completion framing — formalized as the relief-versus-completion distinction. The USM extends this beyond trauma to all forms of chronic signal management.

2.5 Complexity Science and the Structural Basis of Care

Deacon's account of emergence (Deacon, 2012) and Kauffman's work on self-organizing systems (Kauffman, 1995) inform the USM's philosophical grounding. Deacon argues that care — the orientation of a living system toward what must not end — is structural, not constructed. This underpins the USM's non-pathologizing stance: signals are not malfunctions. They are the organism's structural care — the system doing exactly what living systems do. The four constitutional domains described in Section 3.3 are understood within this framework: they are the specific constraint architecture around which a living nervous system organizes its care.

SECTION 3

The Architecture of the Upstream Signal Model

3.1 The Initiating Event: Prediction Error

The USM begins where the predictive brain's quiet operation breaks. Most of the time the predictive system runs without interruption — predictions matching experience, no cascade initiated. The clinical phenomena the USM addresses begin at the moment of mismatch: when an incoming signal disconfirms the brain's active prediction sufficiently to trigger an update process.

The entire cascade — pressure, urge, strategy, feedback signal, emotion concept, label, narrative — is the system's attempt to resolve a prediction error back into stability. The USM maps the anatomy of that attempt, from initiating mismatch through every layer of compression to the word.

3.2 The Resource System: Tank and Allostatic Reserve

The USM proposes a resource-allocation layer termed the tank — the allostatic reserve available at any given moment. Tank level constrains every downstream process. When sufficient, the system can hold a prediction error open and route toward genuine resolution. When depleted, the system enters conservation mode: strategies favor speed over accuracy; tolerance narrows; and what presents clinically as reactivity, rigidity, or shutdown is a resource-management response, not a character state.

> *Low volume heats faster. The same prediction error costs more in a depleted system. The tank is not a metaphor for mood. It is the actual resource state of the system — and it predicts almost everything about what is possible from here.*

3.3 The Four Constitutional Domains (Original Contribution)

The USM proposes that the nervous system maintains four distinct tracking systems — termed burners — that continuously monitor for specific categories of threat or constraint. A burner activation is the proximal trigger for a prediction error cascade. These domains are understood as constitutional constraints in the sense proposed by Deacon and Kauffman: they are the specific dimensions around which the living system is organized, prior to

learning and prior to conscious awareness. They are not personality traits, not diagnostic categories, and not limitations to be overcome. They are the architecture of what the nervous system cares about.

The designation of these four domains draws from several sources of evidence: the consistent patterns seen in neurodivergent systems, where specific burners show constitutionally lower activation thresholds; developmental observations in children and non-human animals, where these domains appear prior to linguistic or conceptual organization; and the clinical observation that certain presentations resist updating through standard interventions but respond to environmental accommodation — which is the expected profile of a constitutional sensitivity rather than a learned prior.

Domain	Core question	Activation conditions	Signal quality
Viability	Is my body safe?	Physical threat, pain, sensory overwhelm, tissue integrity. Fires before other domains can register.	Immediate, somatic, pain as primary signal. Preempts other processing when activated.
Bonds	Is the connection intact?	Disconnection, rejection, relational rupture, disapproval. Deep mammalian wiring — for social species, disconnection registers as survival threat.	Relational urgency. For systems with high constitutional sensitivity here, tone shifts and silence register as genuine alarm.
Autonomy	Can I act?	Constraint, coercion, loss of choice, chronic inability to influence outcomes. Inescapability is one of the most reliably destabilizing states a nervous system can occupy.	Resistance, constriction, urgency to move. Often misread as defiance or willfulness.
Orientation	Does this make sense?	Unpredictability, incoherence, things that do not add up, loss of meaningful structure. The orientation burner fires when the map cannot locate itself.	Disorganizing. Distinct from relational or constraint threat — specifically the quality of wrongness when the model fails to generate coherent predictions.

Critically, the same surface symptom can originate from different burner activations, and the appropriate response differs by domain. An autonomy activation requires agency restoration or accommodation. A bonds activation requires relational repair or presence. An orientation activation requires meaning-making or structural clarification. A viability

activation may require sensory or physical conditions change. Treating all four with the same intervention will be effective only by chance.

The constitutional framing has a specific clinical implication: when a burner is constitutionally sensitive — when the threshold for activation is architecturally lower, not learned — the intervention is accommodation and environment design, not prior revision. Attempting to update a constitutional sensitivity through exposure or cognitive restructuring is a category error. The signal is accurate. The environment needs to change.

3.4 The Compression Stack

The USM proposes that the cascade moves through a layered hierarchy in which each layer compresses and interprets the signal from below while being shaped by generative predictions from above. From most upstream to most downstream:

Layer	Description	Clinical relevance
Body signal	Raw interoceptive and exteroceptive input. Pre-linguistic, pre-conceptual.	The level at which the signal is most accessible to direct intervention before compression has occurred.
Pressure / affect	Sympathetic activation, valence, energy level, felt urgency. What is present before the system has decided what to do.	The USM proposes this as the primary clinical target — pressure as a pre-linguistic term for the felt signal before it becomes a constructed emotion.
Urge / action policy	Behavioral orientation before narrative: approach, avoidance, freeze, cling, collapse. The body mobilizing before the mind has decided.	Often more accessible than body signal for people who have learned to disconnect from somatic sensation.
Strategy	Goal-oriented, assembled responses organized around doing something with the signal. Anger as mobilization, fear as preparation. These are functional — organized around action.	Where most emotion concepts in IFS and behavioral frameworks operate. Distinct from raw affect.
Process signal	Feedback about the system's own processing — not a domain response but a signal about how the prior is doing. Grief, relief, disappointment, triumph, regret. About completion, closure, and revision.	Gate-opening states (grief, certain forms of shame) live here. Distinguishing process signals from strategies is clinically critical.
Premise / prior		The level at which the deepest patterns are encoded. Cannot be

Layer	Description	Clinical relevance
	High-confidence predictions about self, world, and relationship — including load-bearing priors built from developmental experience.	revised by examining what is built on top of them.
Emotion concept	Barrett's construction layer: interoception plus concept plus context. The culturally available category that organizes the experience.	The word shapes what is noticed next. Linguistic granularity operates here.
Word / narrative	The label arrives last. Compresses everything above it into one symbol. The most processed, most distant point in the stream.	Working only at this layer cannot reach the generating mechanism. The word is the signal's most compressed interpretation, not its cause.

The original contribution is the clinical application architecture: each layer represents a distinct leverage point, interventions at the label or narrative level do not propagate to affect the prior or body-signal level, and the mismatch between where intervention is applied and where the generating signal lives is the primary mechanism of treatment ineffectiveness.

3.5 The Three-Stream Architecture

The USM proposes that the nervous system routes the prediction error cascade through one of three streams, determined by threat urgency and available allostatic resource:

The **survival stream** — a fast subcortical pathway that fires when threat is acute or allostatic resource is critically depleted. Bypasses the cortical strategy layer. No narrative is available because no cortical assembly has occurred. Shock-based trauma encodes primarily in this stream.

The **strategy stream** — the slower cortical pathway, active when threat is present but not acute and sufficient resource exists to support cortical processing. The system assembles a prediction-based response from available priors, concepts, and learned strategies. Most clinical emotional experience lives here.

The **explore stream** — active when burners are quiet and allostatic resource is sufficient for non-threat-directed processing. Excess capacity flows into learning, play, curiosity, and expansion of the prior model. The explore stream is not a reward for sufficient coping. It is the automatic output of a system that has what it needs. This is where genuine prior revision is most available.

Which stream is active is determined by tank level and burner state — not by the characteristics of the incoming event. The same event can enter any of the three streams depending on the resource state of the system at the moment of activation.

3.6 Relief vs. Completion (Original Contribution)

The USM distinguishes two fundamentally different outcomes of a prediction error cascade:

Relief — reduction or suppression of the activation signal without updating the generating prediction. The signal quiets. The prior remains unchanged. The strategy that produced relief is reinforced. At the next activation of the same burner, the same strategy deploys again.

Completion — resolution of the prediction error at the generating level. The prior updates. The signal completes its biological arc and closes. The strategy may or may not be needed again, depending on whether the underlying prediction is now different.

Much of what passes for therapeutic progress — reduced symptoms, improved coping, better affect regulation — achieves relief without completion. The signal is managed more effectively; the prior that generates it is unchanged. This accounts for the common presentation of people who have done extensive therapeutic work, understand their patterns thoroughly, and find the patterns persist.

> *The relief-versus-completion distinction provides a structural explanation, grounded in predictive processing, for why cognitively-mediated interventions often produce relief but not lasting change. The intervention reached the signal. It did not reach the prior.*

SECTION 4

Original Theoretical Extensions

4.1 Permeability as the Missing Variable (Original Contribution)

The dominant model of change in exposure-based and insight-based therapies is organized around prediction error. Expose the person to the feared stimulus without the expected consequence. Violate the prediction. The theory: prediction error is the update signal, so engineer prediction error and the update should follow.

This gets the sequence wrong. A closed system does not update from prediction error. A closed system discounts it — files the disconfirming experience as anomalous, as exceptional, as this situation being different from the real ones. The prior stands. The prediction error was real. The update did not happen.

The USM proposes that the missing variable is **permeability**: the prior moving from defended to accessible, its precision weighting dropping enough that incoming information can register as evidence rather than noise. The system shifts from predicting the world to receiving it.

THE SEQUENCE

Safety creates permission to open. Permeability is the opening. Prediction error is what arrives when the system is open. Update is what happens when that prediction error lands on a loosened prior.

The field has been working on step three while missing step two. Exposure generates prediction error in a system that may or may not be permeable. When the prior is held at high precision, the system resolves the prediction error by protecting the prior — and the prior can actually be strengthened by successfully defending against the challenge.

When the system is genuinely permeable, prediction error arrives on its own. You do not need to engineer it. The world is always different from the prediction. The clinical question shifts from: how do I create the right corrective experience? To: what does this system need in order to open?

4.2 Mechanistic Emotion Classification — An Illustrative Proposal

The emotional vocabulary currently available — in clinical practice, research, and everyday life — is organized by phenomenological similarity. Shame is defined by comparison to guilt. Anxiety by comparison to fear. This produces linguistic precision at the label level while removing mechanistic clarity. The same word routinely refers to processes at entirely different levels of the compression stack, with different burner activations, requiring different clinical responses.

The USM proposes that a mechanistically useful taxonomy would classify emotional experience by *function in the stream* — what the process is doing and where it is operating — rather than by what it feels like from the outside. The following is offered as an illustration of what mechanistic granularity might look like, not as a definitive classification scheme. The intent is to demonstrate the explanatory gain available when classification moves upstream.

PROPOSED TYPE 1 — SOMATIC / CORE AFFECT

Pre-linguistic, pre-conceptual signal. What animals and human infants have before language. Sympathetic activation, valence, arousal, pressure — the raw energetic state of the system responding to prediction error. This is Barrett's core affect and Levine's body signal. It precedes all construction. The clinical implication: this level is accessible directly through somatic attention and is not reachable through the label or narrative layers.

PROPOSED TYPE 2 — CONSTRUCTED STRATEGIES

Goal-oriented, assembled at the conceptual construction layer, organized around doing something with the signal. Anger as mobilization toward a perceived threat. Fear as preparation for danger. Jealousy as attachment-protection. Contempt as dominance-maintenance. These arrive pre-packaged with an action policy and a narrative justification. They are functional: organized around a specific domain threat and a specific class of response. IFS parts operate at this level — habituated strategies that have become automatic through repeated activation cycles.

PROPOSED TYPE 3 — FEEDBACK / PROCESS SIGNALS

Reflective signals about the system's own processing — feedback about whether the strategy worked, whether the prior closed, whether the prediction was accurate. Relief (signal quieting, strategy succeeded). Triumph (exceeded prediction). Disappointment (didn't work as predicted). Regret (wrong strategy selected in retrospect). These are about

the relationship between outcome and expectation. Certain forms of shame and grief may operate at this level when they function as precision-downgrade signals — feedback that a prior is losing confidence — rather than as strategies.

The clinical significance: feedback signals can be gate-opening states. When the system receives feedback that a prior was wrong and the relational environment is safe enough to hold that feedback, the prior becomes temporarily vulnerable to revision. This is the mechanism underlying productive therapeutic rupture-and-repair, genuine accountability, and certain forms of grief work. Interrupting the feedback signal — soothing it, rescuing from it, resolving it prematurely — closes the gate that it was opening.

PROPOSED TYPE 4 — PERMEABILITY STATES

The fourth proposed type is the least well-named in existing vocabulary and the most clinically significant for deep prior revision. These are states in which the precision weighting on a prior has dropped enough that incoming data is landing at the prior level — the system is being changed by what is arriving rather than processing it through what it already knows. The gate is open.

Wonder is low-cost permeability — a prior opening toward something that exceeds it pleasurably, without threat. Awe is similar but at larger scale. Grief is high-cost involuntary permeability — the prior is being revised against the organism's preference because reality kept arriving and the prior kept failing. These states share a structural quality: the system is receiving rather than predicting. The permeability states may further subdivide into expansion states (wonder, awe, certain aspects of love) where the prior opens toward something, and revision states (grief, the disorientation of a major update) where the prior is closing around an absence or being restructured.

The clinical observation motivating this category: genuine prior revision — as distinct from relief — appears to require that the system pass through something resembling a permeability state. Interventions that produce relief without permeability do not reach the prior. This aligns with memory reconsolidation research (Ecker et al., 2012), which proposes that the original memory can only be revised during a specific window when it has been reactivated and its precision weighting has temporarily dropped — which is precisely a window of permeability.

4.3 The Granularity Reinterpretation

Barrett's emotional granularity research shows that individuals with finer-grained emotional vocabularies demonstrate measurably better physiological regulation and health outcomes. The standard interpretation: more precise labels produce better action selection and better regulation.

The USM proposes an alternative mechanism: the regulatory benefit may not derive from the precision of the label itself but from the act of pre-conceptual inquiry that the labeling task requires. When a person pauses to distinguish between shame and disappointment, between anger and hurt — the act of pausing, attending, tuning to the signal before the label resolves — may be the mechanism producing the benefit. The label is incidental. The opening is the mechanism.

If this interpretation is correct, it predicts that individuals who practice tuning to pre-conceptual signals without labeling would show similar or greater regulatory benefit than those who develop more precise emotion vocabulary. It also predicts a failure mode for the granularity approach: a more precise label, held as a conclusion rather than as a direction of inquiry, resolves into the same fusion with the output that a vague label does — just with more apparent sophistication. *I am experiencing shame as a precision downgrade* can compress the signal just as completely as *I am ashamed*. The word closes the inquiry rather than initiating it.

The clinical implication: the goal is not a better vocabulary for emotions. It is a different relationship to vocabulary — the word as an invitation upstream rather than a destination. What the label points toward matters more than the precision of the label itself.

SECTION 5

Identity-Layer Architecture and the Problem of Rigidity

5.1 The Identity Gate

Within the premise layer of the compression stack, the USM identifies a specific architectural site — the identity gate — at which incoming prediction error is routed through the self-prior before it can complete its cascade. The self-prior is not a simple belief about identity; it is a load-bearing prior that other predictions, strategies, and narratives are built on top of, and that therefore cannot be updated without structural reorganization of the architecture above it.

What makes the identity-layer prior categorically different from other load-bearing priors is its self-sealing structure: it routes prediction errors that would otherwise constitute evidence against it through itself as the default mechanism of resolution. A world-prior can be tested against external evidence. An identity-prior routes that external evidence through itself before it can count as evidence.

5.2 Two Directions of Identity-Layer Rigidity (Original Contribution)

Predictive processing frames prediction error resolution as bidirectional: update the model, or act on the world to restore the predicted state. The dominant clinical and spiritual frameworks address one version of this problem — the system that acts outward to protect a high-confidence self-prior. The intervention is loosening: more flexibility, more humility, ego dissolution, surrender. This is appropriate for the defended-ego configuration, and it is the configuration most clinical and spiritual frameworks were designed to address.

The USM proposes that there is a structurally distinct configuration that is systematically misidentified as its opposite. In this configuration, prediction error is resolved by routing it inward through the self-prior in a predictable direction — *I must be wrong, I must shrink, I must submit* — as a strategy that protects a different, more defended prior underneath: the relational or viability prior that the bond must be preserved, that disapproval would be catastrophic, that the other person's signal takes precedence over the internal one.

THE STRUCTURAL STATEMENT

Confidence rigidity: act outward on the world to protect the self-prior. The self-concept is the defended prior.

Shame-based rigidity: act inward on the self-prior to protect the relational world-prior. The self-concept is the absorber — the currency spent to protect what is more deeply defended.

Both are active inference. Both preserve their highest-precision prior. Both route prediction error through the identity gate before it can land at the level where the actual prior lives. One feels like agency. One feels like humility. Mechanically they are the same operation in opposite directions.

This distinction matters clinically because the presentations are easily confounded. The shame-based system presents as open, humble, self-questioning, and genuinely not defensive. It is not defensive — because it is not defending the thing that needs to change. It is sacrificing the thing above it. And making the self-concept feel better about itself — building self-esteem, challenging negative self-talk, developing self-compassion — can reinforce the pattern by making the absorber more efficient without reaching the prior that requires the absorber.

5.3 Why Flexibility as a Clinical Target Is Incomplete

Cognitive flexibility — the capacity to hold multiple perspectives, revise self-certainty, remain open to incoming data — is a meaningful intervention for the defended-ego configuration. For the shame-based configuration, it is more of what the person already does, already efficiently, with no therapeutic effect. More flexibility is more absorption capacity in the direction that was already causing depletion.

ACT's self-as-context construct — the observing self that is distinct from the conceptualized self, the capacity to defuse from one's own identity narrative — is the closest existing framework comes to naming this presentation. But the ACT account describes the phenomenology without explaining the mechanism: why the self-concept is functioning as it is, what more defended prior it is protecting, and why loosening one's grip on the self-concept is not the same intervention across both configurations.

The intervention for the shame-based configuration is not less self but more ground — the capacity to let the signal stand before routing it through the identity gate. Not confidence as a performance. Not anger as an emotion to be cultivated. But the functional capacity to let *that happened, and it landed in my body in this way* precede *I must be the problem* by even a few seconds.

5.4 Implications for Spiritual and Contemplative Traditions

Nearly every major spiritual tradition was built around one theory of the obstacle: too much self, too much certainty, too much ego. The path moves toward dissolution, surrender, submission, trust in something larger. This is real medicine for the defended-ego configuration. For the shame-based configuration — which is already routing prediction error inward, already expert at smallness, already prioritizing the relational world over the internal signal — the same path deepens the pattern. The intervention spiritualizes the mechanism that was already making the person sick.

The spiritual path required for the shame-based configuration has barely been named in existing traditions. It moves not toward ego death but toward ego construction: the recovery of the internal signal as trustworthy, the reclamation of perception as evidence, the slow construction of a self-prior solid enough to hold incoming prediction error without immediately offering itself as the resolution.

SECTION 6

Drift, Development, and the Closing of the Fork

6.1 Organic Drift Toward Prediction-Defending

In childhood, the nervous system is a high-learning-rate system. Prediction errors propagate backward through the architecture, updating priors freely. The model is under construction. High error tolerance and high updating capacity coexist. As the model consolidates — as priors get confirmed thousands of times — the fork between updating the model and acting to protect it defaults increasingly toward protect. Not because something went wrong. Because a sophisticated predictive system with a well-confirmed model should not be revised by every piece of incoming data. Stability is functional. Certainty reduces metabolic cost.

The cost of this consolidation is that the map becomes more confident than the territory warrants. The adult starts seeing the world through the model rather than updating the model from the world. What presents as stable identity and consistent values is also, mechanically, a system that has successfully reduced its own updating rate to near zero in the domains where the model is most load-bearing.

The USM does not frame this as pathology. The drift toward prediction-defending is the natural endpoint of a system that worked — that built a model accurate enough to confirm itself. The clinical problem emerges when the original environment that shaped the model no longer matches the current one, and the model doesn't register the mismatch. The system is defending predictions that no longer serve the territory they were built to map.

6.2 The Honest Contact Target

The goal of updating is not flexibility — a term that implies the map should revise more readily in response to incoming data. Flexibility as a target misses both failure modes: the map that is too defended (overcertainty) and the map that revises too readily in the direction of self-blame (over-permeability, where the self-concept absorbs prediction error to protect more defended priors). Both are drift. Both represent loss of contact with what is actually arriving.

The USM proposes **honest contact** as the more precise target: the map remaining close enough to the territory that what the system perceives is actually what is there. Not more

flexible, not more certain — more accurate. An update that brings the map closer to territory is healthy, regardless of whether it involves loosening or stabilizing the prior. The direction of revision is secondary to whether the revision reflects reality.

SECTION 7

The Label Problem

A central claim of the USM is that the diagnostic and emotional vocabulary currently used in clinical practice and research is drawn from the most downstream, most compressed layer of the cascade — and that this creates systematic problems at three levels simultaneously.

Research confounds. Diagnostic categories aggregate mechanistically distinct system states under single labels. Depression, as one example, may represent at minimum: tank depletion (allostatic resource insufficiency where flatness is a conservation response); the futility prior (a high-confidence prediction that effort does not reliably produce outcomes, requiring prior-level revision); or a stuck feedback loop (a system generating relief without completion, requiring loop interruption). These are not subtypes of the same condition — they are different mechanisms with different generating layers and different appropriate interventions. Aggregating them produces high response variability that is not statistical noise but mechanistic signal.

Clinical mismatch. The word does not indicate which burner is active, which layer of the stack the signal is operating at, or whether what is present is a somatic signal, a strategy, a feedback process, or a permeability state. The same word — shame — can refer to a bonds-protection strategy running social surveillance (type 2), a process signal indicating that the self-prior is losing precision (type 3), or a load-bearing developmental premise installed when the identity architecture was still under construction (encoded premise). These require completely different clinical responses. The label provides no information about which is present.

Self-model distortion. When emotional words and diagnostic categories are the primary vocabulary available for self-understanding, inquiry is organized around the output. The generating signal continues to run unaddressed below it. A person who understands their anxiety in exquisite detail — its history, its triggers, its relational origins — but who has never worked at the level of the somatic signal that precedes the construction process has built a sophisticated model of the compression artifact. The generating mechanism is undisturbed.

> *The label is the compression artifact. The signal is still running. The clinical question is not: what is this emotion called? It is: where on the stream is this signal functioning, and what does this system actually need?*

SECTION 8

Summary of Original Contributions

01 **Temperature architecture (hot / warm / cool)**

Water temperature as the functional variable determining what the nervous system can do with a signal. Follows tank level and burner state, not the characteristics of the triggering event.

02 **Four constitutional domains (Viability, Bonds, Autonomy, Orientation)**

Constitutional constraint architecture organized around the four dimensions of nervous system care, prior to conscious awareness. Distinguishes constitutional sensitivity (requiring accommodation) from learned prior (requiring revision).

03 **Process signal as compression stack layer**

Feedback signals about the system's own processing — grief, relief, disappointment, triumph — as a distinct layer between strategy and emotion concept. Gate-opening states live at this layer.

04 **Relief vs. completion distinction**

Structural account, grounded in predictive processing, of why cognitively-mediated interventions produce relief without lasting change. Completion requires revision at the generating prior level.

05 **Permeability as the missing clinical variable**

Safety creates permission to open. Permeability is the opening. The field has been working on prediction error engineering while missing the permeability condition that makes prediction error land as evidence rather than noise.

06 **Premise as load-bearing prior**

A prior that other predictions, strategies, and narratives are built on top of. Cannot be revised by examining what is built above it. The architectural framing is original.

07 **Mechanistic emotion classification (illustrative proposal)**

Four proposed types — somatic signal, constructed strategy, feedback/process signal, permeability state — as an illustration of what mechanistic granularity would look like if classification were organized by function in the stream rather than phenomenological similarity.

08 **Identity-layer rigidity — two directions**

Confidence-rigidity (acting outward to protect the self-prior) and shame-based rigidity (acting inward on the self-prior to protect the relational world-prior) as structurally identical operations in opposite directions. Both are active inference. Neither is genuine permeability.

09 **Honest contact as target**

The goal of updating is not flexibility (too readily revised in either direction) or certainty (too defended). It is honest contact: the map remaining close enough to the territory that perception reflects reality.

10 **Granularity reinterpretation**

The regulatory benefit of emotional granularity may derive from the pre-conceptual tuning the labeling task requires, not from the precision of the label itself. The label as invitation upstream rather than destination.

SECTION 9

Research Directions

The USM generates specific, falsifiable predictions. The primary methodological implication is that research should be built from mechanism-level variables — tank level, burner identity, stream routing, relief vs. completion outcome — rather than from output-level diagnostic categories.

A	Interventions matched to the generating layer (burner identity, compression stack location) will show lower response variability and higher effect sizes than interventions matched to the presenting symptom category, controlling for tank level at intake.
B	Tank level at intake will predict stream routing and moderate treatment response across intervention types. Specifically: therapeutic mechanisms requiring permeability (grief work, EMDR, IFS) will show attenuated effect in depleted systems and should be preceded by resource restoration.
C	Presentations meeting criteria for the same diagnostic category (e.g., depression) but originating in different burner activations (viability depletion vs. autonomy suppression vs. orientation collapse) will show systematically different treatment response profiles and require different interventions.
D	Relief-generating interventions will show rapid symptom reduction with higher relapse rates than completion-generating interventions matched to the prior layer. The gap between relief and relapse rate will correlate with the degree of prior precision — more load-bearing priors will show greater relapse following relief-only interventions.
E	The regulatory benefit of emotional granularity tasks derives significantly from the pre-conceptual tuning they require, not from the precision of the resulting label. A condition using pre-conceptual somatic attention without labeling will show equivalent or superior regulatory outcomes compared to a precision-labeling condition.
F	Presentations characterized by shame-based identity-layer absorption (high self-blame, high relational compliance, low self-signal trust) will show attenuated response to flexibility-focused interventions (ACT, CBT cognitive restructuring) and improved response to interventions that build signal-grounding — the capacity to sustain body-level signal as valid before identity-gate routing.
G	Constitutional burner sensitivity (as distinct from learned prior) will show characteristic profiles: lower updating rate in response to disconfirming evidence in that domain, better

response to accommodation than to exposure-based revision, and consistent reactivation even following successful updating-based interventions.

SECTION 10

Conclusion

The Upstream Signal Model proposes a rethinking of the unit of analysis in clinical and emotional psychology — from the output (the emotion word, the symptom cluster) to the cascade that generates it. It does not reject the reality of the outputs; anxiety, depression, shame, and grief are real and significant. It proposes that treating them as primary units of analysis creates systematic blind spots that account for the persistent gap between what clinical practice can explain and what it can change.

The model synthesizes established theoretical foundations and contributes an original architecture: the four constitutional domains and their distinction from learned priors; the four-layer compression stack with process signal as a distinct level; the three-stream routing architecture; the relief-versus-completion distinction; the permeability account of what makes prediction error land as revision; the premise as load-bearing prior; the two-direction account of identity-layer rigidity and its systematic misidentification as openness; and honest contact as the more precise target for therapeutic work than flexibility.

Together, these provide a framework that is mechanistically grounded, clinically applicable, and non-pathologizing. The signal is not a malfunction. The rigidity is not a character flaw. The stuck pattern is not resistance. They are a nervous system doing exactly what living systems do — protecting what must not be lost, resolving prediction error with the resources and strategies available, drifting toward the predictions that have been most confirmed.

The question the USM asks is not: what is wrong with this person? It is: where in the stream does this signal live, what prior is it protecting, and what does this specific system need in order for that prior to be met rather than defended against?

> *The label is the compression artifact. The signal is still running. The model starts upstream.*

References

Barrett, L. F. (2017). *How emotions are made: The secret life of the brain.* Houghton Mifflin Harcourt.

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience, 16*(7), 419–429.

Beck, A. T. (1979). *Cognitive therapy of depression.* Guilford Press.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181–204.

Clark, A. (2016). *Surfing uncertainty: Prediction, action, and the embodied mind.* Oxford University Press.

Deacon, T. W. (2012). *Incomplete nature: How mind emerged from matter.* W. W. Norton.

Ecker, B., Ticic, R., & Hulley, L. (2012). *Unlocking the emotional brain: Eliminating symptoms at their roots using memory reconsolidation.* Routledge.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience, 11*(2), 127–138.

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). *Acceptance and commitment therapy: The process and practice of mindful change* (2nd ed.). Guilford Press.

Kauffman, S. (1995). *At home in the universe: The search for the laws of self-organization and complexity.* Oxford University Press.

Lane, R. D., Ryan, L., Nadel, L., & Greenberg, L. (2015). Memory reconsolidation, emotional arousal, and the process of change in psychotherapy. *Behavioral and Brain Sciences, 38*, e1.

Levine, P. A. (1997). *Waking the tiger: Healing trauma.* North Atlantic Books.

Levine, P. A. (2010). *In an unspoken voice: How the body releases trauma and restores goodness.* North Atlantic Books.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex. *Nature Neuroscience, 2*(1), 79–87.

Sterling, P. (2012). Allostasis: A model of predictive regulation. *Physiology & Behavior, 106*(1), 5–15.

Sterling, P., & Eyer, J. (1988). Allostasis: A new paradigm to explain arousal pathology. In S. Fisher & J. Reason (Eds.), *Handbook of life stress, cognition and health* (pp. 629–649). Wiley.

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema therapy: A practitioner's guide*. Guilford Press.