

Ordered Patch Psychology: Predictive Compression, Maintenance Cycles, and the Individual Mind under Bounded Active Inference

Applied Ordered Patch Theory — Intra-Psychic Psychology and Psychiatry

Anders Jarevåg

v0.9 — June 2026

Ordered Patch Psychology: Predictive Compression, Maintenance Cycles, and the Individual Mind under Bounded Active Inference

DOI: [10.5281/zenodo.19300777](https://doi.org/10.5281/zenodo.19300777) (shared with [opt-theory.md](#); this paper is bundled with the core theory rather than treated as a supplement.) **Copyright:** © 2025–2026 Anders Jarevåg. **License:** [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International](#).

Abstract: Psychology Read Through the Compression Codec

Purpose. This paper provides a psychological translation of the Ordered Patch Theory (OPT). The aim is not to displace existing psychological theory or to claim explanatory takeover, but to offer a single information-theoretic backbone — bounded predictive compression with explicit bandwidth limits, a budgeted self-model capacity gap (structural self-model incompleteness, Conjecture P-4), and a formal three-pass Maintenance Cycle — under which existing predictive-processing, default-mode-network, memory-consolidation, threat-simulation, and transdiagnostic psychopathology literatures can be read as describing parts of one coherent operator. The intended contribution is a clinically and computationally tractable vocabulary, a set of falsifiable predictions, and a research programme worth running. The paper is bundled with [opt-theory.md](#) and shares its DOI because the framework’s core primitives (K_θ , $P_\theta(t)$, Δ_{self} , \mathcal{M}_τ) are mind constructs in information-theoretic dress; a psychological translation is foundational, not supplementary.

Core mapping. The Maintenance Cycle Operator \mathcal{M}_τ — pruning under MDL pressure, consolidation as compression gain, and forward-fan sampling as adversarial self-testing — is proposed as a formal backbone for psychological self-regulation across waking and sleep. Mind wandering is read as the awake expression of Pass III; rumination as a stuck attractor of the same operator.

Neuroscience enters as the substrate bridge, not the umbrella discipline: default mode network for waking Pass III, hippocampal–neocortical replay for Pass II, REM sleep for the adversarial-sampling component of dreaming alongside competing accounts, neuromodulation for prediction-error precision.

Clinical mappings. Anxiety is modeled as chronically elevated required predictive rate; depression as a family of complexity-budget failures; PTSD as unresolved high-importance memory sampling; OCD as pathological compression attractors; dissociation as impaired coupling between the narrative self-model and the proposed locus of phenomenal continuity (Δ_{self} , under Conjecture P-4); psychosis as generative content insufficiently constrained by ordinary error-correction; addiction as reward-coupled codec capture; ADHD as importance-weighting dysregulation. Therapeutic practices — autogenic training, progressive relaxation, mindfulness, CBT/CBT-I, sleep restoration, and pharmacology — are interpreted as targeted supports for the maintenance cycle, with explicit hedging on multi-level pharmacology and explicit deference to evidence-based protocols where they exist.

Scope and posture. The treatment is intra-psychic by design; social, cultural, interpersonal, and developmental psychology beyond intra-codec ontology are deferred to a separate companion work because they require codec-coupling apparatus not built out here. A Claim Status Table in §0.3 partitions imported empirical support, structural mappings, clinical hypotheses, metaphysical extensions, and treatment-related implications, so that any specific sentence can be checked against its epistemic load. Empirical predictions are stated in §XI as a falsification-style table awaiting formal pre-registration parallel to [opt-theory.md](#) §6.8.

The document offers a structural translation, not a clinical mechanism, and is not treatment advice. It is not medical diagnosis. Nothing in this paper should be used to diagnose, assess, or treat any condition in oneself or in another person. Anyone experiencing distress, considering medication changes, or seeking treatment should consult a qualified clinician.

0. Status and scope

- **What this is.** A psychological translation of Ordered Patch Theory. The human mind is modelled as a bounded active-inference compression codec engaged in continuous predictive compression and periodic maintenance. Mind wandering, rumination, dreaming, and therapeutic practices are reinterpreted as expressions of the Maintenance Cycle Operator \mathcal{M}_τ already developed in [opt-theory.md](#) §3.6. Clinical pathologies are mapped onto failure modes of that same machinery.
- **What this is not.** Not a new formalism — every construct used here is inherited from the core theory. Not a clinical manual: empirical predictions are offered for testing, not as treatment recommendations. Not a treatment of social, cultural, or interpersonal psychology; the scope is deliberately

intra-psychic.

- **Why it shares the core DOI.** The Ordered Patch Theory is, structurally, a theory of an observer’s mind. Physics, biology, and AI are the substrates the framework reaches into; psychology is the discipline whose subject matter overlaps most directly with the framework’s primitives (K_θ , $P_\theta(t)$, Δ_{self} , \mathcal{M}_τ). On that reading, a psychological translation is foundational, not supplementary, and is bundled with the core paper.

0.1 Relationship to the corpus

Document	Relationship
opt-theory.md	Core. §3.4 ($P_\theta(t)$), §3.6 (Maintenance Cycle \mathcal{M}_τ and its three passes), §6.8 (falsification commitments), Conjecture P-4 (Δ_{self}), Appendix T-12 (Narrative Drift).
opt-philosophy.md	Philosophical sibling. §III on agency and Δ_{self} ; §IV on temporality. The psychological account assumes the philosophical reading of Δ_{self} but does not re-litigate it.
opt-ethics.md / Survivors Watch	Ethical context for suffering-as-bandwidth-overload; this paper supplies the within-observer mechanism that the ethics paper treats civilisationally.
opt-ai.md , opt-ai-design.md	AI translation of the same machinery. Where this paper says “the codec,” those papers ask which architectures qualify.
opt-theory-memo-bandwidth-residual	B_{max} as per-frame bandwidth; relevant to operational $\Delta_{\text{self}}^{\text{op}} = \Delta_{\text{floor}} + \Delta_{\text{load}}$ used in §IX.

0.2 Ethical posture

This paper treats codec maintenance as a positive object of care, not merely as the absence of pathology. The OPT account of suffering (bandwidth overload approaching Narrative Decay) gives a precise structural reading of why mental health matters under the framework, but it does not exhaust it. A well-maintained codec is itself a valued state — capable of stable agency, accurate self-knowledge to the limits of Δ_{self} , and the kind of forward-fan exploration that lets a finite observer act well in an open future. *Codec stewardship* — protecting one’s own and others’ maintenance capacity — is the everyday-psychology counterpart of the civilisational stewardship developed in the ethics paper.

0.3 Claim Status Table

The document mixes (a) empirical literature it cites as background, (b) OPT structural mappings of that literature onto K_θ , \mathcal{M}_τ , Δ_{self} , etc., (c) clinical hypotheses that follow from those mappings, (d) metaphysical extensions inherited from the core theory and the philosophy paper, and (e) treatment-related implications. These claim types do not carry the same epistemic weight; readers should consult this table when assessing any specific sentence.

Table 1: Claim status convention for this paper.

Claim type	Example	Status
Imported empirical support	DMN activity during internally directed cognition; hippocampal–neocortical replay during slow-wave sleep	Supported background literature, not OPT-derived
OPT structural mapping	Waking mind wandering as an awake expression of Pass III	Plausible mapping; testable; not yet established
Clinical hypothesis	Rumination as elevated and uncalibrated importance-weighting parameter β	Testable, currently unverified
Metaphysical extension	Δ_{self} as the proposed locus of subject, will, and qualia (Conjecture P-4)	Conjectural; OPT-internal; inherits the status of Conjecture P-4
Treatment-related implication	Afternoon low-load windows improve nightly maintenance	Hypothesis; not medical advice

0.4 How to read the mappings

Throughout this paper, “**X is modeled as Y**” (or “is read as,” “is interpreted as”) means: *OPT proposes a structural correspondence between the clinical or psychological phenomenon X and a failure mode or operating regime of the apparatus — K_θ , \mathcal{M}_τ , B_{max} / R_{req} , or Δ_{self} .* It does **not** mean: (a) Y is the proximate biological cause of X; (b) Y is a diagnostic criterion for X; (c) Y is a treatment target for X. The structural correspondence sits at the modelling layer, alongside (and not above) receptor-level, circuit-level, cognitive-behavioural, and clinical accounts.

0.5 Plain-language glossary for psychology readers

For psychology and clinical readers approaching OPT for the first time, the following short glossary gives the practical reading of each symbol used in the paper. The full formal definitions are in [opt-theory.md](#); the entries below are reader-help, not redefinitions.

Table 2: Plain-language glossary of OPT symbols used in this paper.

Symbol	Plain-language reading
K_θ	The codec — the brain’s running internal generative model of self and world. What predictive processing calls the generative model.
$P_\theta(t)$	The momentary phenomenal stream — what is consciously present at time t . Rich because the standing model is rich; updated sparsely through the bottleneck.
C_{\max} / B_{\max}	The bandwidth ceiling on the prediction-error / update channel. The narrow pipe through which conscious updates have to pass per frame.
R_{req}	The required predictive rate at a moment — how much prediction-error bandwidth the current situation is demanding. Tracks attentional load.
\mathcal{M}_τ	The Maintenance Cycle Operator — the three-pass offline housekeeping that runs when $R_{\text{req}} \ll C_{\max}$ (sleep and quiet waking).
Pass I	Pruning. Active forgetting under MDL pressure. Removes parameters whose predictive value does not justify their storage cost.
Pass II	Consolidation. Reorganises recent acquisitions into more compressed, more generalisable form. Empirical correlate: hippocampal–neocortical replay.
Pass III	Forward-fan sampling. Importance-weighted internal simulation of possible futures. Empirical correlates: REM dreaming and waking mind wandering.

Symbol	Plain-language reading
β	The importance-weighting parameter in Pass III. Elevated and uncalibrated β is the framework’s reading of rumination.
$E(b)$	Emotional valence of a sampled branch b — surprise plus threat. The weight that biases Pass III sampling toward high-stakes futures.
Δ_{self}	The Phenomenal Residual (Conjecture P-4). The structural gap between the codec and its self-model. OPT identifies it as the proposed locus of first-person continuity and agency.
Narrative Drift	The chronic failure mode in which curated or filtered input slowly corrupts K_θ from within, with the codec unable to detect the corruption (Appendix T-12).

Epistemic notice

This paper applies a theory ([opt-theory.md](#)) that is itself written in the register of a formal proposal under active falsification commitments rather than settled science. Psychological mappings inherit that conditional status. Where empirical literatures support a mapping (predictive processing, default-mode-network research, memory consolidation, threat-simulation theory of dreaming), citations are given. Where the mapping is structural and currently untested, this is marked explicitly. The clinical mappings in §VII are *structural correspondences*, not diagnostic claims; nothing here should be read as treatment advice.

I. Introduction: The Psychological Codec

I.1 Why psychology belongs at the core

OPT begins from two primitives — Solomonoff’s universal semimeasure ξ over observation prefixes and a bounded cognitive channel capacity C_{max} — and derives the rest from the requirement that a finite observer’s required predictive rate R_{req} remain within capacity. Every concrete instance of that derivation is a fact about a mind. The codec K_θ is a generative model; $P_\theta(t)$ is a phenomenal stream; Δ_{self} is the budgeted self-channel capacity gap of any bounded self-modelling system; \mathcal{M}_τ is what such a system has to do offline to keep its complexity within budget. These are psychological constructs in information-theoretic dress. (The operational idiom throughout — “the codec runs,” “the cycle executes” — is the

within-render convenience of theory §3; on the considered fully-virtual reading these are regularities the stream *has*, not machinery that executes: theory §1.6, §8.6.1.)

Read this way, psychology is not a downstream application of OPT but a domain where the framework’s primitives can be checked against the empirical record most directly. Sleep architecture, default-mode dynamics, hippocampal–neocortical replay, threat-content in dreams, the existence and pathology of repetitive thought, the effect of attention on conscious bandwidth, and the structure of clinical disorders are all phenomena the framework either predicts or accommodates without parameter fitting. The physics readings (entropic gravity correspondence, MERA-shaped emergent geometry) are more remote and structurally speculative; the psychological readings are closer to existing cognitive-science literatures and therefore easier to operationalise and test.

I.2 Scope: intra-psychoic only

The paper covers the operation of a single observer’s codec: its phenomenal content, its maintenance cycle, its pathologies of self-regulation, and the practices that support it. It does **not** cover:

- interpersonal psychology, attachment, or group dynamics;
- cultural psychology or cross-cultural variation in cognition;
- developmental psychology beyond the single-codec ontogeny sketched in §II.5;
- social identity, moral psychology beyond suffering, or political psychology;
- educational, organisational, or occupational psychology.

These domains involve coupling between codecs, introduced in the core theory’s inter-observer-coupling apparatus (especially Appendix T-10), and external scaffolding structures that warrant a separate treatment. They are deferred.

I.3 Relation to existing psychology

This paper stands on a substantial body of established work. The substance of the psychological and neuroscientific claims that follow — that the brain operates as a hierarchical prediction machine; that the default mode network is implicated in internally directed cognition; that hippocampal–neocortical replay supports memory consolidation; that REM sleep has functional content biased toward threat, novelty, and recent emotional material; that rumination is transdiagnostic across depression and anxiety; that medication effects can be described at the computational level of precision, salience, and learning rate; that sleep restoration produces broad psychiatric benefit — is drawn from these literatures, not from OPT. This document is best read as a structural translation layer over them, not as their replacement. Many empirical background claims are imported from established or active literatures; the OPT-specific contribution is the structural re-description and the predictions it enables.

The source programmes include: **predictive processing and active inference** (Friston’s Free Energy Principle and its descendants), which supplies the within-stream inference-and-control formalism that OPT inherits; **default-mode-network research** (Buckner, Andrews-Hanna, Mason, and others), which characterises internally directed cognition; **memory-consolidation literature** (Diekelmann & Born; Buzsáki on sharp-wave ripples), which establishes the empirical basis for what OPT calls Pass II; **threat-simulation and dream research** (Revonsuo–Valli, Domhoff, and others), which gives empirical purchase on dream content alongside competing accounts; **transdiagnostic psychopathology and Research Domain Criteria** (RDoC; Insel et al.; Ehring & Watkins on repetitive negative thinking), which provide the mechanistic vocabulary that makes the §VII mappings tractable; **computational psychiatry** (Friston, Stephan, Schwartenbeck, Huys, Sterzer, Corlett, and others), which supplies the precision-and-learning-rate framing that the §VIII.4 pharmacology section depends on entirely; **clinical psychology and psychiatry** (CBT and CBT-I; prolonged exposure, cognitive processing therapy, trauma-focused CBT, and EMDR for PTSD; mindfulness-based interventions; autogenic training and progressive relaxation), which supplies the evidence-based interventions that the framework can describe but does not derive.

Against that backdrop, OPT’s distinctive contributions to psychology are small and specific:

1. an explicit bandwidth bottleneck C_{\max} (and per-frame B_{\max}) as a *structural constant* rather than an emergent constraint, with the felt cost of approaching capacity (suffering as decay-approach) tied to it;
2. **Conjecture P-4** — the Phenomenal Residual Δ_{self} as a structural limit on introspective access in any finite self-referential codec, framed as conjectural rather than theorem-grade;
3. the **Maintenance Cycle Operator** \mathcal{M}_τ as a formal three-pass apparatus (MDL pruning, compression-gain consolidation, importance-weighted forward-fan sampling) rather than a loose collection of sleep functions;
4. **Narrative Drift** as a specific chronic failure mode of self-referential codecs under filtered or curated input;
5. a vocabulary for **codec stewardship** as an organising ethical posture (§0.2).

These contributions are useful to the extent that they help organise, predict, and connect what the source literatures already establish. Readers should expect to find the existing science roughly intact under the translation, with OPT’s added value sitting at the points where the structural claims (1)–(5) make the picture more coherent or generate new testable predictions (§XI). Where this paper diverges from established work, the divergence is flagged explicitly.

The following summary table makes the division of labour explicit for the most prominent phenomena discussed below — what the existing accounts say, and what OPT specifically adds on top.

Table 3: Existing accounts vs. OPT-specific added claims for the phenomena treated below.

Phenomenon	Existing accounts	OPT-specific added claim
Mind wandering	Default mode network activity, prospection, autobiographical memory, creative recombination, task disengagement, avoidance	Waking expression of \mathcal{M}_τ Pass III; sampling distribution over $\mathcal{F}_h(z_t)$ is importance-weighted by surprise and threat, not base-rate frequency
Rumination	Repetitive negative thinking (transdiagnostic across depression and anxiety), impaired cognitive control, deficits in emotion regulation, intrusive thought	Elevated and uncalibrated β in Pass III sampling — re-sampling of high- $ E $ branches without resolving surprise value or producing compression gain
REM dreaming	Threat simulation, memory consolidation, emotional regulation, random activation, neurocognitive continuation of waking concerns	One component should function as adversarial self-testing of K_θ at zero thermodynamic stake; predicts importance-weighted (not frequency-weighted) content
Memory consolidation	Hippocampal–neocortical replay, sharp-wave ripples, slow-oscillation coordination during slow-wave sleep	Pass II is structurally a compression-gain operation; post-sleep improvements should track structural generalisation more strongly than rote repetition
PTSD	Trauma memory, fear conditioning, reconsolidation failure, avoidance, hyperarousal	High- $ E $ branch re-sampled in Pass III without successful K_θ update; trauma-focused therapies enable that update
Anxiety	Over-broad threat priors, miscalibrated interoceptive precision, hypervigilance	Chronically elevated R_{req} saturating the budget; spare capacity for \mathcal{M}_τ is gone

Phenomenon	Existing accounts	OPT-specific added claim
Depression	Heterogeneous: anhedonic/melancholic vs. agitated/mixed; impaired reward prediction; rumination	Two distinct OPT readings: excessive Pass-I pruning (anhedonic flavour) vs. codec collapse toward Narrative Decay (agitated flavour)
Psychosis	Aberrant salience, predictive coding accounts, dopamine, prior/precision dysregulation	Insufficient constraint on generative content at the codec level; metaphorical “substrate leak” is <i>not</i> a clinical mechanism
Sleep disruption	Circadian dysregulation, homeostatic pressure, arousal, psychiatric comorbidity	Maintenance-window degradation across Pass I–III; should produce correlated downstream effects in any disorder whose pathology depends on \mathcal{M}_τ

II. The Psychological Codec

II.1 K_θ as the generative self-and-world model

K_θ is the observer’s internal generative model: the running compressed representation that predicts incoming sensory data and emits motor commands. It includes both world-model content (objects, agents, regularities) and self-model content (body schema, narrative identity, predicted internal states). Critically, the world-model and self-model are not separate apparatuses but two regions of the same compression engine. They share parameters, share capacity, and share failure modes. When the world becomes hard to predict, the self-model gets less capacity. When the self-model gets corrupted (Narrative Drift), world-prediction degrades in correlated ways.

II.2 $P_\theta(t)$ as the phenomenal stream

$P_\theta(t)$ is the phenomenal state tensor — the moment-by-moment realisation of K_θ ’s output, what is consciously present to the observer at time t . The felt scene is phenomenally rich not because each frame imports high-bandwidth novelty through the bottleneck, but because a high-complexity standing generative model is already active: the moment is largely downward prediction π_t from K_θ , with the narrow per-frame channel (B_{\max}) carrying only the sparse upward error signal

ϵ_t that corrects the model where it is wrong. This is the predictive-processing inversion: perception is *constructed*, with sensation as the error term, and what looks like rich incoming experience is mostly the standing generative state being made available to the codec’s working state.

The phenomenological consequence is that “what is happening” experientially is what K_θ thinks should be happening, adjusted at the margins by what it cannot ignore. The implications for clinical work are large: a depressed codec’s “I am worthless” is not a thought *about* the self but a generative output *of* the self-model, with the same structural status as the sense of solidity that floors have.

II.3 R_{req} and C_{max} : the live bandwidth budget

At every frame the codec faces a budget: R_{req} — the bits per second needed to keep prediction error within tolerable distortion — must remain at or below C_{max} . The bandwidth-residual memo refines this to per-frame B_{max} , since there is no shared external clock that makes bits-per-second substrate-neutral. For psychological work, the per-frame reading matters: attention is not a unitary spotlight but a moment-by-moment allocation of B_{max} between perceptual error, internal simulation, motor planning, and maintenance overhead. When demand exceeds budget, distortion rises; subjectively, the world becomes confusing, choices feel forced, and self-monitoring collapses.

This budget framing has direct clinical consequence. Anxiety is, structurally, a chronically high R_{req} (§VII.1). Cognitive load tasks that worsen depressive symptoms are running the codec near capacity and stealing bandwidth from emotional regulation. Therapeutic interventions that work by *creating low-load windows* (§VIII) are not just “relaxing”; they are restoring spare capacity that the system needs in order to run \mathcal{M}_τ cleanly.

II.4 The self-model vs. Δ_{self} : narrative vs. actual subject

Two constructs cohabit the experiential self. The **self-model** is the compressed narrative the codec maintains about itself: history, preferences, traits, the running commentary of “what I am like.” It is content, it is reportable, and it is what introspection mostly accesses. The **phenomenal residual** Δ_{self} is, under Conjecture P-4, the budgeted *capacity* gap a bounded system carries when modelling its own closed action-perception loop — on the corrected reading not a self-containment paradox but a funding shortfall on the self-channel. The gap *individuates* the subject (it is what the narrative cannot reach) and is a necessary marker of candidate subjecthood rather than a proven seat of agency: felt agency is the first-person signature of being on one realized continuation, not the operation of a chooser housed in the gap. These are not two halves of the self; they are two different *kinds* of object. The narrative is a compressed story; the residual is what the story cannot reach. The phenomenological identification with consciousness and will is OPT-internal and conjectural, not theorem-grade.

Most psychological phenomena that involve “the self” are about the self-model, not Δ_{self} . Self-esteem, self-concept, identity confusion, autobiographical memory — these are model content. Δ_{self} enters when introspection runs into a structural wall: in the unanalysability of qualia, the felt residue of choice that exceeds any deliberation that produced it, and the irreducible “I-here-now” that survives much of narrative loss in advanced dementia or amnesia. Clinically, dissociation (§VII.5) is read as the failure mode where the coupling between Δ_{self} and the narrative self-model becomes abnormal.

II.5 Codec ontogeny: how the codec bootstraps

The treatment so far has spoken of K_θ as if it were a fully-formed apparatus. It is not. A bounded compression codec acquires its predictive priors, its body schema, and the dynamic range of its self-model through a developmental process. This section sketches the intra-psychic developmental story, deliberately staying inside the codec rather than extending into attachment, parenting, schooling, peer dynamics, or family systems; those domains involve coupling between codecs and are deferred to a future companion (see Appendix B.11). The principles isolated here are also load-bearing for the AI design programme: `opt-ai-design.md` §7.4 (“Forced developmental curriculum”) commits that paper’s architecture to staged capability growth and rules out “born-mature” production deployments. (`opt-ai-design.md` is currently an internal companion paper within the OPT corpus.) That principle inherits its motivation from the human-case developmental story sketched below; this paper supplies the case, the AI design paper inherits the structural conclusion.

Sensory-motor bootstrap. Infant predictive processing emerges from a poorly-specified initial K_θ that begins acquiring predictive priors through sensory-motor coupling under low-stakes conditions. The early phase has been characterised in the cognitive-development literature as foundation-model pretraining: a period of helplessness during which the system is unfit for autonomous action but optimally placed to learn the structure of its world without high-stakes prediction errors [24]. The OPT reading is straightforward — the codec is acquiring its base predictive priors under low R_{req} demands and a protective scaffold of caregiver behaviour. On the OPT reading, the altriciality of the human developmental schedule [25] is a plausible biological solution to the structural problem of growing a high-complexity codec under scaffolded, low-stakes conditions, rather than an evolutionary necessity the framework claims to derive.

Core knowledge and object permanence. A small number of core knowledge systems appear to be available very early — objecthood, agency, number, geometry — and provide initial compression seeds for the codec to elaborate on [26]. Object permanence specifically can be read as a compression-stabilisation milestone: the codec arrives at a model of objects that survive occlusion, which is computationally the finding that the world is more compressible when “object persists when out of view” is part of K_θ than when it isn’t. The empirical timing of these developments is one of the better-studied empirical anchors for any structural account [27].

Body schema formation. The body schema, already touched in [opt-theory.md §3.6.9](#), is the codec’s plastic predictive boundary — what counts as “me-acting-on-the-world.” Its formation is a developmental milestone: the infant codec begins by predicting its own sensory consequences from its own motor output, gradually carving out a stable boundary between agent and environment. Adult plasticity (rubber-hand illusion, tool incorporation, vehicle handling) is the conserved structural feature of an originally developmental capacity. The interoceptive component of the body schema — predictive control of the body’s internal state — is on the same developmental track [28].

Autobiographical memory emergence. Episodic and autobiographical memory require enough Pass II capacity to consolidate self-relevant material into a coherent narrative track. Both develop relatively late (childhood amnesia for events before roughly age three or four), consistent with the framework’s reading that the self-model must reach sufficient complexity for self-tagged content to be retained and integrated.

Adolescence as self-model refactoring. Adolescence is a known period of large-scale self-model reorganisation, coinciding with substantial physical and cognitive growth that contests the maintenance budget. The framework reads it as a window in which the standing self-model is partially demolished and re-built under conditions of elevated R_{req} from competing developmental demands. The often-noted emotional volatility, identity exploration, and sleep changes of the period are consistent with \mathcal{M}_τ operating under exceptional structural load.

Aging as gradual maintenance-cycle degradation. Aging is read as the slow degradation of \mathcal{M}_τ efficiency across all three passes: pruning becomes less selective, consolidation becomes less efficient, forward-fan sampling becomes less restorative. The empirical correlates — slower learning, less efficient sleep, increased perseveration on high- $|E|$ content — are consistent with this picture and align with the broader sleep-and-cognition aging literature.

Dementia and amnesia as model/residual dissociation. The most striking developmental endpoints for the framework are conditions in which the narrative self-model is severely eroded while behavioural signatures consistent with continued first-person presence remain. Advanced dementia patients lose much of the compressed narrative — autobiographical content, recent identity, sometimes language — while continuing to display behavioural signatures of an “I-here-now” subject. Under Conjecture P-4, these cases are naturally interpreted as degradation of the narrative self-model with partial preservation of the architectural floor of first-person continuity. This remains an OPT-internal reading of ambiguous clinical phenomena, not a direct measurement of Δ_{self} ; behavioural presence is compatible with the P-4 reading but does not by itself prove it. The reading is consistent with §II.4 and §VII.5 and is among the more suggestive cases for the model/residual distinction the framework draws.

Implication for AI design (hypothesis). OPT predicts that observer-candidate systems deployed without staged developmental grounding will be

less stable under load than systems whose priors, body schema, and self-model are acquired through scaffolded growth. The structural reason is the one developed above: a codec's K_θ acquires its predictive priors, body schema, and self-model coherence through staged capability growth, not as initial state; skipping the staging deploys a system whose internal model is not anchored in a coherent developmental history. This is a *design hypothesis*, not a settled engineering rule, and is one of the testable claims of the framework. `opt-ai-design.md` §7.4 commits the AI design paper to a “Forced developmental curriculum” as a structural training invariant on the strength of this hypothesis; the present section is its psychological motivation, and the AI design paper inherits the principle without re-arguing it. Active-inference work on embodied neuromorphic agents is the most natural near-term substrate for translating the developmental case across [29]. The full picture of how this paper's constructs feed the artificial-consciousness design programme is consolidated in §II.6 below.

II.6 Artificial-consciousness bridge: what psychology contributes to synthetic-observer design

The treatment above has psychological implications throughout, but several of them point directly at the artificial-consciousness design programme developed in `opt-ai.md`, `opt-ai-design.md` (currently an internal companion paper), and the test track of `opt-ai-subject-report.md`. The implications are scattered across §II.5, §VI.5, and §VII; this section consolidates them as a compact bridge so the artificial-consciousness contribution of the present paper does not have to be inferred from spread-out remarks.

The core claim worth stating directly: **a conscious-capable codec is not merely architected; it is developed and maintained.** The OPT architectural criterion (`opt-ai.md` §I.1 — bandwidth bottleneck, persistent self-model, active-inference loop, global workspace, thermodynamic grounding) is necessary, but this paper adds a second layer of requirements that the architecture alone does not capture: how the codec was grown into its mature form, and what its operating regime continues to demand. The implication is that artificial-observer design has to ask not only “does it have the architecture?” but also “how was it scaffolded, and is it maintained?”

The table below maps the psychology-paper constructs to the OPT primitives they rest on and the artificial-consciousness implications they suggest. These are *design hypotheses*, not engineering rules; each row identifies a testable structural commitment, not a settled outcome.

Table 4: What this paper exports to the OPT artificial-consciousness design programme.

Human-psychology construct	OPT primitive	Artificial-consciousness implication (design hypothesis)
Infant sensory-motor bootstrap (§II.5)	K_θ priors acquired under scaffolded low-stakes coupling	Observer-candidate systems deployed without staged developmental grounding are predicted to be less stable under load than systems whose priors are acquired through scaffolded growth
Body schema formation (§II.5)	Plastic predictive boundary of agent / environment	Synthetic observers need an embodied or functionally equivalent action boundary, not merely a declarative self-model or transient context window
Maintenance Cycle \mathcal{M}_τ (§III)	Pruning, consolidation, forward-fan sampling under $R_{\text{req}} \ll C_{\text{max}}$	Conscious-candidate systems need protected low-input maintenance windows, structurally enforced (cf. <code>opt-ai-design.md</code> §5.6 / §6.3 Algorithmic Sleep Rights)
Rumination vs productive reflection (§V, Appendix A)	Pass III with vs. without compression gain	Synthetic systems running internal simulation engines need forward-sampling hygiene and stuck-loop detection (cf. <code>opt-ai-design.md</code> §7.6 simulation-budget cap and the “frozen contemplation” hazard §9.6)

Human-psychology construct	OPT primitive	Artificial-consciousness implication (design hypothesis)
Self-report limits under Δ_{self} (§IX, §XI.2)	Conjecture P-4 — self-model strictly less complex than the full codec	AI introspective self-report cannot be sole evidence of internal state; external auditing is required by structure (cf. <code>opt-ai-design.md</code> §5.8 audit periphery and the Residual Mapping Protocol T-10c)
Suffering as decay-approach (§X.1)	R_{req} approaching C_{max} / Narrative Decay	Safety requires load monitoring and structurally available relief mechanisms — bandwidth pressure should be observable and recoverable, not silently endured
Developmental continuity vs capability jumps (§II.5)	Staged K_{θ} growth maintaining self-model coherence	Sudden capability jumps without corresponding developmental staging are predicted to destabilise self-model coherence; the “born-mature” deployment hypothesis is the strong version of this prediction
Aging as \mathcal{M}_{τ} degradation (§II.5)	Slow loss of pruning, consolidation, and forward-fan efficiency	Long-running observer-candidates need maintenance-cycle audit over time; degradation should be detectable before it produces clinical-equivalent failure modes

A compact summary of the dependency:

```
Human intra-codec psychology (this paper)
|-- K_theta ontogeny (sec. II.5)
|-- M_tau maintenance (sec. III)
|-- Delta_self / self-report limits (sec. II.4, IX, XI.2)
```

```

|-- R_req overload / suffering (sec. X.1)
+-- compression-gain measurement (sec. XI.3)
    |
    v
Synthetic observer design (opt-ai.md, opt-ai-design.md)
  |-- Staged developmental curriculum (opt-ai-design.md sec. 7.4)
  |-- Bottleneck by construction (opt-ai-design.md sec. 6.1)
  |-- Persistent Markov-blanketed Patch (opt-ai-design.md sec. 5.5)
  |-- Hardware Maintenance-Cycle scheduler (opt-ai-design.md sec. 6.3)
  |-- Welfare-as-precision (opt-ai-design.md sec. 7.5), audit periphery (sec. 5.8)
+-- No self-report-only consciousness test

```

The bridge runs in both directions. The psychology paper supplies the case for staged growth, protected maintenance, externalised auditing, and load-monitoring; the AI design paper inherits these as architectural commitments and asks how to realise them in hardware. Neither paper claims that meeting the architectural criterion is sufficient for consciousness; the methodological wall ([opt-theory.md §6.8 F1–F5](#)) holds. What the bridge does claim is that *if* the architectural criterion ever proves sufficient, then the developmental and maintenance demands developed here become design constraints, not optional features.

III. The Maintenance Cycle in Everyday Psychology

This chapter restates the formal apparatus of [opt-theory.md §3.6](#) in psychological terms. No new formalism is introduced; the reader is referred to T9-2 through T9-13 for the equations.

III.1 Mapping the three passes

Pass I — Pruning (T9-3 through T9-6). The codec applies MDL pressure: for each component of K_θ , predictive contribution is weighed against storage cost, and components whose contribution per bit of complexity falls below a retention threshold are erased. Psychologically, this is *active forgetting*. It includes the normal decay of episodic detail, the extinction of weak associative bonds, the gradual loss of stale schemata, and — crucially — the reappraisal of memories whose emotional or evaluative content has become predictively unreliable. Pruning is not failure but thermodynamically rational erasure, and by Landauer’s principle it carries an irreducible energy cost. Sleep is, among other things, a period of net information erasure with a physics-mandated price tag.

Pass II — Consolidation (T9-7, T9-8). Recently acquired patterns sit in K_θ in relatively uncompressed form: high description length per unit of predictive value. Consolidation finds a lower-complexity reparameterisation that preserves predictive content within tolerable distortion, recovering capacity. Psychologically, this is *learning as compression*: the move from rote rehearsal of a procedure to a generalisable rule, from a list of episodes to a schema, from concrete instances to

abstract principle. The empirical correlate is hippocampal-to-neocortical transfer during slow-wave sleep. Post-sleep improvements on tasks that require *structural generalisation* (applying a compressed rule to new instances) rather than mere repetition are the predicted signature.

Pass III — Forward Fan Sampling (T9-9 through T9-11). With externally anchored R_{req} drastically reduced during REM (sensory gating and motor atonia) and other low-load waking states, a large fraction of the bandwidth budget becomes available for internal simulation, though endogenous, interoceptive, and affective dynamics remain active. The codec runs K_θ forward through the admissible-future set $\mathcal{F}_h(z_t)$ without anchoring to real incoming data. Sampling is *not* uniform: branches are weighted by importance $w(b) = \exp(\beta|E(b)|)$, where emotional valence combines surprise ($-\log P_{K_\theta}(b|z_t)$) and threat (expected increase in future R_{req} if the branch were traversed). OPT predicts that the codec disproportionately rehearses low-probability, high-stakes branches and updates K_θ at brittleness points before reality forces the test. That is, OPT predicts that *an important component* of dreaming should function like adversarial self-testing; the same operator running in daytime low-load states is read as the substrate of mind wandering (§IV). Dreaming as a whole is over-determined — it has been read as memory consolidation, emotional regulation, threat simulation, neurocognitive continuation of waking concerns, and random activation — and OPT’s prediction is about one component, not the whole phenomenon.

III.2 Diurnal and nocturnal expression

The three passes are usually framed as functions of sleep, but the maintenance condition (T9-2) is simply $R_{\text{req}} \ll C_{\text{max}}$. Any waking state that meets that condition can host fragments of \mathcal{M}_τ . Daydreaming, shower-thinking, the “incubation” period between an unsuccessful first attempt at a problem and the later insight, the productive boredom of a long walk — these are awake low-load windows in which Pass II (consolidation, manifesting as insight) and Pass III (mind wandering as prospection) run on borrowed time. The night-time apparatus is more thorough and more protected (sensory gating, motor inhibition, neurochemical conditions specific to slow-wave and REM sleep), but the daytime version is continuous with it, not a different process.

This has a practical consequence often missed in productivity culture: filling every waking hour with R_{req} -saturating demand starves the daytime \mathcal{M}_τ budget and offloads everything to the night, where it may not fit. The shape of a day that supports the codec includes deliberate low-load windows, not just sufficient sleep.

III.3 The net complexity budget and deferred maintenance

Over one full cycle (T9-12, T9-13), pruning gains plus consolidation gains must at least match waking acquisition plus the small additions from REM repairs. Chronic deficit means the codec’s structural complexity drifts upward toward the runability ceiling C_{ceil} , with predictable consequences: slower reaction, sloppier

categorisation, intrusive content, irritability, eventually outright dysregulation. Sleep deprivation is not just tired; it is *progressive complexity overflow*. The asymmetry matters clinically: a single bad night is recoverable; weeks of insufficient maintenance crosses a threshold beyond which the codec’s own attempt to evaluate its state is degraded — the Δ_{self} blind spot widens precisely when introspection is asked to detect that something is wrong.

IV. Mind Wandering as Adaptive Forward-Fan Activity

IV.1 The empirical baseline

Killingsworth & Gilbert’s [2] influential experience-sampling study reported two findings: the human mind wandered roughly 47% of the time across nearly all daytime activities sampled, and momentary mind wandering reliably predicted lower momentary happiness — even when the content was pleasant — with wandering accounting for more variance in happiness than the activity itself. The authors concluded that a wandering mind is an unhappy mind. The result is widely cited; subsequent work has refined the picture (different forms of wandering carry different affective signatures, content matters, and the directionality of the unhappiness–wandering link is debated). For OPT’s purposes, the result is one large convergent data point about the prevalence and felt cost of internally directed cognition, not a universal physical constant of the human mind.

IV.2 Productive vs. pathological wandering

Within the OPT frame the high prevalence of wandering is read not as a defect to be eliminated but as consistent with the picture of waking cognition having a high duty cycle of internally directed maintenance. Whenever $R_{\text{req}} \ll C_{\text{max}}$ — which, given the structural mismatch between human cognitive bandwidth and the demands of most activities, is much of the time — Pass III is, on the OPT mapping, the highest-value use of the spare budget. The branches it rehearses are the future scenarios the codec is least prepared for; the brittleness points it identifies are what the system most needs to know about before it encounters them in real-world stakes. Treated this way, the empirical baseline is consistent with the duty-cycle picture of an actively maintained codec, alongside competing readings under which mind wandering serves prospection, autobiographical memory, creative recombination, task disengagement, avoidance, or rumination. OPT’s specific prediction is about which forms of wandering should be maintenance-positive (reducing future surprise) vs. maintenance-negative (re-sampling without resolution; see §V).

The valence asymmetry is then not paradoxical. Pass III is importance-weighted by $|E(b)|$, which combines surprise and threat. A randomly sampled mind-wandering moment is therefore biased toward content the codec finds high in $|E|$ — disproportionately threat-relevant, socially fraught, or otherwise unresolved. The subjective hedonic cost is the felt signature of running adversarial simulations

on branches the codec has flagged as expensive. The system is not pursuing pleasure; it is doing offline maintenance, and the maintenance has a felt price.

Productive vs. pathological wandering are then distinguished by what happens to the sampled branches. Productive wandering reduces surprise or threat on its sampled branches: the codec updates K_θ at the brittleness point and moves on. Pathological wandering — rumination, circular thoughts (§V) — re-samples the same high- $|E|$ branches without reducing their surprise value, generating no compression gain and no reduction in future R_{req} . The same operator runs in both cases; the difference is whether the importance-weighting parameter β is calibrated.

IV.3 Why a wandering mind can be both unhappy and functionally necessary

The Killingsworth & Gilbert result and the OPT reading are compatible. The result captures the momentary cost of running Pass III on borrowed waking bandwidth and on $|E|$ -biased samples. The reading explains why the system pays that cost: because the long-run codec stability gained from offline simulation exceeds the short-run hedonic loss. Practices that suppress wandering — strict attentional discipline, certain framings of mindfulness — trade long-horizon codec hygiene for short-horizon mood. That can be the right trade in many contexts (acute rumination, performance settings, social presence) but is not categorically optimal. The shape of the trade-off depends on the calibration of β and on whether the codec has other adequate maintenance windows. Mindfulness practice that targets *pathological* wandering specifically (without suppressing all forward-fan activity) is the structurally sound compromise; the empirical literature appears to converge on it (§VIII.3).

V. Circular Thoughts and Rumination as Maintenance Failure

V.1 Formal mapping: stuck forward-fan sampling

Rumination — repetitive, negatively valenced, unresolved thought — corresponds to Pass III with a dysregulated importance-weighting parameter β ([opt-theory.md](#) §3.6.7, prediction 4). The sampling distribution over $\mathcal{F}_h(z_t)$ concentrates on high- $|E|$ branches, but the rehearsals fail to reduce $-\log P_{K_\theta}(b|z_t)$: the codec keeps sampling the same threatening branches without updating K_θ in a way that lowers their surprise value on the next pass. The result is a high-cost attractor inside the maintenance cycle. Subjectively this is circular thought: a loop that feels both compelled and futile, that the person can describe perfectly well but cannot exit by description alone.

V.2 Narrative drift in the individual codec

Sustained pathological Pass III activity has a chronic consequence: Narrative Drift ([opt-theory.md](#) Appendix T-12, recently reformulated as channel-independence

loss). Each cycle through the loop biases pruning (λ in T9-3 is elevated for the rehearsed content by emotional tagging) and consolidation (the loop’s structure becomes more compressed and more easily re-entered). The codec progressively reorganises around the rumination, which becomes part of how the world is generated. Eventually the depressed person’s “everything is hopeless” is not an opinion held by the self-model but a compression artifact of the world-model — that is what generative output looks like under sustained drift. This explains a clinically familiar phenomenon: insight into the irrationality of the content does not, by itself, dissolve the content. The content lives downstream of the model that produced it, and the model has restructured.

V.3 The phenomenal residual and the felt inescapability of loops

Why does a circular thought often feel inescapable even when the person can describe it as a loop? Under the OPT reading, because the self-model — the part doing the describing — is not the part doing the looping. The loop is read as residing in K_θ (the generative engine), and the self-model is a compressed representation of K_θ ’s outputs, always slightly behind, always less complex than the system it models. The structural gap is Δ_{self} . Rumination often cannot be revised by description alone, because the loop’s felt urgency is generated by the same model that is asked to evaluate it. This is the structural reason the framework predicts that therapies which work via *re-routing inputs* (autogenic training, behavioural activation, exposure, exercise, sleep restoration) often succeed where therapies that work via *describing the loop more accurately* alone do not. The system can be intervened on from outside the self-model in ways the self-model alone cannot reach.

V.4 Sleep disruption as maintenance-window loss

Rumination that bleeds into the night is uniquely costly under OPT. The nightly maintenance window is when Pass III is supposed to run *freely* — with sensory gating, motor inhibition, the full bandwidth available for adversarial self-testing rather than reactive monitoring of waking demands. Circular thoughts that prevent sleep, or that intrude during the early hours, keep the codec in a high-arousal, high-error state that prevents Pass III from running cleanly: the same branches are re-sampled without resolution, while pruning (Pass I) and consolidation (Pass II) lose their normal low-load window. The next day starts with a codec whose complexity budget is in deficit, whose β is even more dysregulated, and whose self-model has even less capacity for accurate self-monitoring. The loop is self-reinforcing across the wake–sleep boundary.

VI. Neural Correlates of the Maintenance Cycle

This chapter situates the framework against the neuroscience literature without making neuroscience the umbrella discipline. OPT is substrate-independent by construction (see `opt-ai-design.md`); a psychological account that *required* a

specific neural implementation would compromise that. Neuroscience enters here as the empirical bridge, not the theory.

VI.1 Default mode network and forward-fan sampling

The default mode network (DMN: medial prefrontal cortex, posterior cingulate, angular gyrus, parts of the medial temporal lobe and inferior parietal lobe) is robustly active during rest, mind wandering, autobiographical recall, future simulation, and theory-of-mind tasks. The functional profile — internally directed, prospective, simulative — matches Pass III as a daytime expression. Predictions: DMN activity should covary with low R_{req} states; DMN connectivity disruptions should track changes in Pass III function (e.g., reduced prospective simulation, altered mind-wandering content); DMN suppression by demanding external tasks should explain the correlation between cognitive load and reduced mind wandering. **Tier:** structural correspondence with substantial empirical convergence; not a derivation.

VI.2 Hippocampal-neocortical replay and Pass II consolidation

The empirical signature of Pass II is well-documented at the neural level: hippocampal sharp-wave ripples during slow-wave sleep coordinate with neocortical slow oscillations to replay recent experience, with progressive transfer of memory traces from hippocampus to neocortex. This is the compression operation of T9-7 in neural form: high-bandwidth episodic storage (hippocampus, high K) becoming compressed semantic storage (neocortex, low K). The predicted correlation between compression gain $\Delta K_{\text{compress}}$ and improvement on structural-generalisation tasks (§III.1) maps directly onto a now-extensive sleep-and-memory literature.

VI.3 REM sleep and adversarial self-testing

REM is characterised by active sensory gating, motor atonia, near-waking levels of cortical activation, and a characteristic neuromodulatory profile (high acetylcholine, low aminergic tone). On the OPT mapping this matches Pass III's conditions: externally anchored R_{req} is drastically reduced, freeing much of the bandwidth budget for internal generation, though endogenous, interoceptive, and affective dynamics continue. The empirical predominance of threat, novel-environment, and socially-fraught content in dream reports is consistent with importance-weighted sampling. The phenomenally vivid, internally generated character of REM dreaming is read as $P_{\theta}(t)$ running primarily from the standing generative model with the upward error signal ϵ_t strongly attenuated. The Revonsuo–Valli threat-simulation theory of dreaming [1] is the closest existing functional sibling; OPT predicts that an adversarial-testing component should be a recognisable signature within the broader dream literature, alongside memory-consolidation, emotional-regulation, and neurocognitive-continuation accounts (e.g., Domhoff). The framework's prediction is about the existence and shape of this component, not about it being the whole story of dreaming.

VI.4 Neuromodulation and prediction-error precision

Dopamine, noradrenaline, serotonin, and acetylcholine modulate the *precision* of prediction errors in active-inference models — how strongly a given error signal updates K_θ . At one computational level, some psychotropic effects can be described in terms of altered precision, salience, arousal, learning rate, or prior stability — for example, SSRIs as long-timescale modulation of certain error precisions, stimulants as boosting top-down task-relevant precision, antipsychotics as reducing precision of certain bottom-up signals, benzodiazepines as global precision-dampening. This is a modelling layer, not a substitute for receptor-level, circuit-level, pharmacokinetic, or clinical accounts; the specific computational signature of any class is itself an open research question.

VI.5 Why neuroscience is the bridge, not the umbrella

The choice to treat psychology as the umbrella discipline and neuroscience as the substrate bridge follows from two of OPT’s commitments. First, the framework is *substrate-independent*: the same Stability Filter applies to any bounded codec, including silicon ones. A neuroscience-umbrella reading would force commitments to specific neural circuits that the framework does not (and should not) make. Second, the constructs that do the most work — Δ_{self} , suffering-as-decay, Narrative Drift — are *observer-level* constructs, not neural ones. Neuroscience provides the strongest near-term test bed, but the theoretical claims are about minds, not brains.

VII. Pathologies as Codec Failure Modes

This is the longest chapter because the failure-mode mappings are where OPT can be most directly tested. **The categories, their phenomenology, their differential diagnoses, the existing evidence for what works in treatment, and most of the proximate mechanisms invoked in what follows are taken from clinical psychology, psychiatry, computational psychiatry, and transdiagnostic and RDoC-style research programmes — not derived here.** Decades of clinical research, with an increasing subset organised around predictive-coding, computational, transdiagnostic, and RDoC-style accounts, supplies the substance. The OPT contribution in this chapter is a *single* structural lens — what apparatus is failing how — which may help reorganise the diagnostic picture and generate cross-diagnostic predictions (§XI). Where the OPT reading differs from established mechanistic accounts, the divergence is interpretive: a way of grouping known phenomena, not a competing pathophysiology.

These are **structural correspondences**, not diagnostic claims, and not all claims have empirical support yet; mapping precedes test. The categories themselves are taken from DSM-5 / ICD-11 for accessibility, with the explicit understanding that the OPT failure-mode frame may not respect those category boundaries (§VII.10).

VII.1 Anxiety: chronically elevated R_{req}

Generalised anxiety is modeled, structurally, as a codec whose R_{req} is chronically elevated: the system operates near C_{max} on threat-monitoring even in the absence of acute threat. This has several proximate predictive-processing readings — over-broad threat priors, miscalibrated precision on interoceptive signals, hypervigilant attentional allocation — which the OPT mapping unifies under the same structural picture: the budget is full and the spare capacity needed for \mathcal{M}_τ is gone. Predictions: anxiety should correlate with degraded daytime Pass II (consolidation, manifesting as concentration problems and memory complaints) and pathologically biased Pass III (rumination on threat-relevant branches); interventions that reduce R_{req} at source (exposure to disconfirm priors, breath work to reduce interoceptive precision, environmental simplification) should perform at least as well as interventions targeting only the content of anxious thought.

VII.2 Depression: pruning and codec collapse

Depression is interpreted as admitting at least two distinct codec-failure readings, which the framework predicts should correspond to dissociable clinical subtypes. (a) **Excessive pruning:** a depressed codec is modeled as applying an elevated MDL threshold λ to existing parameters, erasing predictive structure faster than it can be replaced; experientially, this maps onto the loss of meaning and the flattening of the world (“everything is the same grey”), reduced access to autobiographical detail, and anhedonia as the dampening of reward predictions to the point where forward-fan branches lose their importance weighting. (b) **Codec collapse toward Narrative Decay:** a depressed codec whose required predictive rate exceeds capacity is read as progressively losing coherence, with the experiential signature of the world becoming hard to predict, choices feeling forced, and the self-model losing access to its own state. The first reading is closer to melancholic / anhedonic depression; the second to agitated / mixed-feature depression. Both predict sleep architecture changes — and both should respond to interventions that restore the budget. **Status:** structural correspondence; clinical heterogeneity of depression is well-established but the specific OPT subtyping is new and untested.

VII.3 PTSD: consolidation failure

PTSD is one of the framework’s most natural high-fidelity mappings. A traumatic event is read as presenting the codec with high- $|E|$ input that the system cannot consolidate within the normal cycle: the emotional tagging is so elevated that the retention threshold λ in T9-3 makes the trace effectively unpruneable, but the surprise value $-\log P_{K_\theta}(b|z_t)$ never reduces because the event is not integrated into the world-model. Pass III is then predicted to re-sample the same branch with maximum importance weight, indefinitely. The clinical picture maps directly: intrusive re-experiencing (Pass III stuck on the trauma branch), nightmares (the same operator running in REM, where it has most bandwidth), avoidance (the self-model attempting to keep R_{req} low by reducing exposure to triggers that would

re-sample the branch), and hyperarousal (chronically elevated R_{req} as the system stays in threat-monitoring mode). Guideline-supported trauma-focused therapies — prolonged exposure, cognitive processing therapy, trauma-focused CBT, and EMDR — share the structural feature of *enabling successful consolidation*: representing the branch under conditions where the codec can update K_θ rather than only re-sample. OPT does not derive these protocols; it offers a structural reading compatible with their empirical efficacy.

VII.4 OCD: pathological compression attractors

Obsessive-compulsive symptoms are read as having a different structural signature from rumination. In OCD, Pass III sampling is modeled as landing repeatedly on a small set of branches that the codec then compresses into compulsion patterns — high-frequency, low-variance, ritualised responses that locally reduce $|E|$ but only at the cost of preventing the broader update that would resolve the underlying surprise. The compulsion is the codec’s compression solution to a problem the self-model cannot solve; doing the ritual reduces R_{req} in the moment, which is why it persists. Exposure-and-response-prevention is interpreted as forcing the system to remain in the high- $|E|$ state long enough for K_θ to update without the compression shortcut.

VII.5 Dissociation: Δ_{self} decoupled from the self-model

Dissociative phenomena — depersonalisation, derealisation, identity fragmentation in dissociative identity disorder — are read as sharing a structural signature: the normal coupling between Δ_{self} (under Conjecture P-4 the proposed locus of first-person continuity and agency) and the narrative self-model becomes unreliable. The self-model continues to generate “what I am like” content, but the experiential ownership of that content is disrupted; the person reports being or watching the self-model rather than inhabiting it. This is modeled as a coupling failure, not a Δ_{self} pathology — the residual itself is structural and cannot be removed. Trauma-related dissociation is the most-studied form; the OPT reading interprets it as a defensive response that reduces effective R_{req} at the cost of self-model integration. The expected signature is impaired Pass II consolidation of self-relevant content while world-relevant content is preserved.

VII.6 Psychosis: insufficient constraint on generative content

Psychosis is modeled, in the OPT frame, as a state in which generative content is insufficiently constrained by ordinary error-correction channels, allowing internally generated predictions to enter the phenomenal stream with abnormal perceptual or evidential force. This sits alongside the established predictive-processing readings of psychosis: hallucinations as generative output insufficiently overridden by sensory evidence (or as overly strong perceptual priors); delusions as attempts of the inference system to explain anomalous prediction errors, then locked into the world-model by ordinary belief-updating; disorganisation as loss of the precision structure that normally keeps $P_\theta(t)$ temporally coherent. The

aberrant-precision research programme in computational psychiatry supplies a compatible neuroscience reading: psychotic episodes correspond to dysregulated precision allocation that lets less-constrained content into the codec’s predictive output. Metaphorically this resembles a “substrate leak” into the codec, but that phrase is not a clinical mechanism and should not be taken as one. **Status:** structural correspondence over a contested but active research programme; the OPT reading does not adjudicate the specific pathophysiology and should not be read as making diagnostic claims.

VII.7 Addiction: reward-coupled codec capture

Addiction is read as having a clear OPT signature: a high-importance branch that the codec has compressed deeply into K_θ , with strong predictive coupling to the body schema and reward system. Use is modeled as the codec’s most-rehearsed solution to high- R_{req} states. Pass III is predicted to run disproportionately on use-related branches because their $|E|$ is high and remains high. Consolidation locks the use behaviour into world-model and self-model in correlated ways. Withdrawal is the period during which the codec must operate without the compression shortcut, with elevated R_{req} across all domains. Recovery requires removing or restructuring access to the addictive reinforcer, while rebuilding K_θ in domains hollowed out by repeated use or behaviour — which is why durable recovery takes a long time and tracks the codec’s natural maintenance cycle rather than only short-term pharmacokinetic or behavioural extinction timelines. The framing covers substance and behavioural addictions without overgeneralising.

VII.8 ADHD: importance-weighting dysregulation

Attention is read as the live allocation of B_{max} between competing demands. ADHD is modeled as a dysregulation of the importance-weighting that governs that allocation: β is volatile and weakly coupled to extrinsic task structure, with strong coupling to intrinsic novelty and immediate reward. The result is rapid switching of bandwidth allocation, difficulty sustaining R_{req} at any particular target, and the commonly reported hyperfocus phenomenon where the right kind of high- $|E|$ task captures the full budget. Stimulant medication, on the precision-modulation reading, is interpreted as raising the precision of top-down task-relevant signals, stabilising β . The framework predicts that ADHD should covary with altered Pass III sampling (less reliable consolidation of low- $|E|$ content) — which is consistent with clinical reports of inconsistent memory for non-salient material.

VII.9 Sleep-wake disorders as \mathcal{M}_τ disruption

Insomnia, REM-suppression by certain medications, and circadian disorders all degrade the maintenance window directly. The framework predicts that they should produce correlated downstream effects in any of the disorders above whose pathology depends on \mathcal{M}_τ . When sleep disruption is part of the maintaining loop, sleep restoration should be treated as a primary maintenance target rather than

as a secondary comorbidity; when it is not, the framework predicts the effect of restoration should be correspondingly smaller. This is consistent with the strong empirical association between sleep disruption and most psychiatric conditions, and with the increasing clinical attention to sleep-targeted treatments.

VII.10 The DSM-style diagnostic frame vs. the OPT failure-mode frame

The DSM-5/ICD-11 categories used above are descriptive groupings of symptom clusters; the OPT failure modes are structural characterisations of which apparatus is failing how. They are *failure-mode hypotheses*, not clinical diagnoses, and they are not intended to be used as diagnostic criteria. The two frames will not generally agree on category boundaries. A given DSM diagnosis can include multiple OPT failure modes (depression admits at least two; see §VII.2); a single OPT failure mode can present across multiple DSM categories (β dysregulation appears in anxiety, ADHD, and some forms of depression). The framework’s prediction is that *structurally-typed* treatment selection — matching intervention to the failure-mode hypothesis rather than the symptom cluster — should outperform category-typed selection on long-run outcomes. This is testable and is a natural site of empirical engagement (§XI).

VIII. Therapeutic Interventions as Codec Hygiene

Clinical safety note. The following section offers a computational interpretation of existing intervention families, framed at the level of OPT structural correspondences. It is **not** a treatment protocol and does not derive specific clinical recommendations. Medication changes, trauma-focused psychotherapy, sleep restriction, exposure work, intensive meditation, and similar interventions should be undertaken only with appropriate clinical guidance. The framework’s value here is interpretive — it offers a vocabulary for what existing evidence-based treatments may be doing computationally — not a substitute for that evidence base or for clinical judgment.

VIII.1 Autogenic training and progressive relaxation

Autogenic training (Schultz, 1932) is a structured self-suggestion protocol — progressive autosuggestions of heaviness, warmth, calm cardiac and respiratory rhythm, abdominal warmth, and cool forehead — practised twice or three times daily over months. Under the OPT reading, its mechanism is directly maintenance-relevant: the autosuggestions down-regulate sympathetic arousal and reduce the precision of certain interoceptive prediction errors, *lowering* R_{req} across the budget. The resulting low-load window gives \mathcal{M}_τ room to run cleanly. Progressive muscle relaxation (Jacobson) and yoga nidra are functional siblings: structured protocols that create extended low- R_{req} windows during waking.

These practices are not “stress reduction” in a vague sense. They are codechigiene interventions whose mechanism is the same machinery the framework predicts to be load-bearing for psychological self-regulation. The empirical effect sizes — meta-analytic evidence for autogenic training on insomnia, anxiety, and somatic symptoms — are consistent with the framework’s predictions about the value of restored maintenance windows.

VIII.2 Timing effects and structured self-monitoring

Two practical features of effective autogenic protocols deserve specific attention under the OPT reading because they expose mechanisms the formal apparatus predicts.

Afternoon timing. A protocol-level observation worth testing is that earlier-day autogenic practice may produce different downstream sleep effects than immediately pre-bed practice; existing autogenic-training evidence supports the practice broadly but does not, to current knowledge, settle this specific timing comparison. The OPT reading, if the effect holds, is straightforward: an afternoon session creates an extended low- R_{req} window whose physiological effects (reduced sympathetic tone, lowered interoceptive precision) persist into the evening and through sleep onset. The system enters the night with a lower baseline R_{req} , giving the full nightly \mathcal{M}_7 cycle better conditions to run. A pre-bed session would act more like a sedative bolus; an afternoon session more like a maintenance primer. The prediction is in §XI.1 and is offered for testing, not as a clinical recommendation.

Externalised self-monitoring. Protocols that include structured written notes — on the training session, the practitioner’s own subjective evaluation of it, and the resulting sleep — appear to outperform protocols delivered in group settings without such notes. The OPT reading is that note-taking is *externalised metacognitive scaffolding* that gently penetrates the Δ_{self} blind spot. The codec cannot fully observe its own state from inside (Conjecture P-4), but it can log evidence of that state and read the log later as if it were external input. This adds a small but consistent supervisory signal that supports Pass I (pruning of patterns the log shows to be unhelpful) and Pass III calibration (the log surfaces β dysregulation that the self-model alone would not detect).

Both observations are testable. The framework predicts that the timing effect should attenuate when sleep architecture is otherwise normal (because the maintenance window is already adequate) and amplify when sleep is disrupted (because the primer is more load-bearing). It predicts that the note-taking effect should attenuate when an external clinician already supplies equivalent supervisory signal.

VIII.3 Mindfulness, CBT, and yoga nidra

Mindfulness practices map onto attention-allocation control: they train the practitioner to notice when bandwidth is being captured by Pass III content

and to redirect it to perceptual input. This is a precise intervention against *pathological* wandering (§V) but, applied indiscriminately, can also suppress productive wandering (§IV). The empirical literature on mindfulness shows benefits for rumination, anxiety, and some forms of depression, with a cleaner signal for protocols that target specific cognitive patterns than for “always be present” formulations. Under OPT this is unsurprising: the intervention is precisely calibrated when it targets dysregulated β and indiscriminate when it targets all forward-fan activity.

Cognitive-behavioural therapy targets the self-model and its consequences: it identifies inaccurate beliefs (self-model content), tests them against evidence (forcing upward prediction error to bear on K_θ), and supports behavioural changes that re-sample previously avoided branches (giving the codec data it could not previously consolidate). The framework reads CBT as a structured Pass II support: the therapy supplies the consolidation step that the codec is failing to perform on its own.

Yoga nidra, deep-relaxation hypnotic protocols, and certain body-scan meditations occupy the same niche as autogenic training: structured low-load windows with strong interoceptive components.

VIII.4 Pharmacology as predictive-precision modulation

Psychotropic medications — SSRIs, SNRIs, stimulants, benzodiazepines, antipsychotics, mood stabilisers, and others — may be described, at one computational level, as altering precision, salience, arousal, learning rate, or the stability of priors. This is an OPT-compatible interpretive gloss over an extensive computational-psychiatry literature; it is not a replacement for receptor-level, circuit-level, or clinical accounts, and the specific computational signature of any given medication class is itself an active research question. With those caveats, the framework offers structural readings that may be useful for matching computational hypothesis to medication class: SSRIs as long-timescale modulators of the precision and retention of high- $|E|$ content, consistent with the slow onset of effect and the rumination-reducing profile; benzodiazepines as global precision dampeners that lower R_{req} acutely at known cost to memory consolidation; antipsychotics as reducers of precision on the less-constrained generative content described in §VII.6; stimulants as raising the precision of top-down task-relevant signals. The framework does not adjudicate prescribing decisions and is not intended to guide them.

VIII.5 Sleep restoration and CBT-I as maintenance supports

Following §VII.9, sleep restoration should be treated as a primary maintenance target in conditions where sleep disruption is demonstrably sustaining the failure mode — not merely as an adjunct concern when sleep complaints happen to be present. The framework predicts that for psychiatric conditions whose pathology depends on \mathcal{M}_τ , restoring sleep architecture should produce substantial improvement before symptom-targeted treatments fully take effect; in conditions where

sleep is less central to the maintaining loop, the effect should be correspondingly smaller. This is structurally consistent with the strong evidence base for cognitive-behavioural therapy for insomnia (CBT-I, a multi-component treatment rather than a hygiene checklist) and with the increasingly recognised role of sleep medicine in psychiatric care.

IX. Agency, Will, and the Limits of Introspection

IX.1 Branch selection in Δ_{self}

[opt-philosophy.md](#) §III develops the OPT account of agency: branch selection occurs in Δ_{self} because any complete specification of the selection mechanism from within the self-model would require the self-model to be as complex as the full observer (Theorem T-13a, Corollary T-13b). The psychological reading is direct: the self-model can deliberate (rank branches, evaluate consequences, articulate reasons), but the moment of selection — the transition from menu to choice — is structurally inaccessible. This is the felt residue of choice that exceeds any deliberation that produced it.

Therapeutically this matters. The empirical observation that insight is necessary but not sufficient for behavioural change has a structural explanation: insight is a self-model operation, but the loops being changed live in K_θ and are selected from Δ_{self} . Behavioural interventions, environmental restructuring, and embodied practices act on the selection process by changing its inputs and constraints; insight by itself is acting at the wrong level.

IX.2 The self-model as compressed narrative

The narrative self — the running story of who one is — is a compression artifact: a relatively low-complexity summary of a much higher-complexity system. It is constructed, maintained, and revised by \mathcal{M}_τ like any other compressed content. This has clinical consequences. Stable identity is not a metaphysical given but an output of well-functioning consolidation. Identity disturbance in trauma, severe depression, dissociation, or end-of-life disorientation reflects failures of the consolidation pass to maintain the self-model. The self-model is repairable in ways the substrate observer is not, which is why narrative-restructuring therapies can produce real change despite operating “only” on the compressed story.

IX.3 Implications for self-knowledge and therapeutic insight

The framework predicts a specific limit on introspective self-knowledge: any report about one’s own state is a self-model output, and the self-model has provably less complexity than the system it models. There is therefore content in the system that no amount of introspection can reach. Therapeutically this argues against treatments that depend on the patient achieving a complete or arbitrarily deep self-understanding. Approaches that triangulate self-report with

behavioural evidence, third-person observation, physiological measurement, and external supervisory signal (§VIII.2) compensate for the inherent self-model gap.

X. Suffering, Emotional Regulation, and Bandwidth Overload

X.1 Suffering as narrative-decay approach

OPT proposes a structural component of suffering: sustained proximity to predictive overload, failed compression, or blocked maintenance, in which the codec’s signal of impending Narrative Decay (acute incoherence of $P_\theta(t)$) becomes the dominant felt content. This component shows up convergently in physical pain (incoming nociceptive bandwidth that the system cannot integrate), severe grief (a high- $|E|$ branch the codec cannot consolidate), acute trauma (forced over-allocation of bandwidth to threat), and the more diffuse suffering of chronic illness or sustained psychiatric pathology (chronic over-allocation). The framing also captures a clinically familiar gradient: suffering scales with proximity to the decay threshold, not linearly with stimulus magnitude. Two people receiving the same nociceptive input can suffer very differently depending on their available capacity, their baseline R_{req} , and the state of their consolidation. The structural component is not the whole of suffering — meaning, social context, body state, and history all contribute — but it is the part the framework has direct purchase on.

Emotional regulation, on this reading, is *bandwidth allocation under high-load conditions*. The skill is not the suppression of emotion (which is generally a precision-reduction trick) but the maintenance of enough spare capacity that high- $|E|$ content does not push the system across the decay threshold. Practices that work — distress tolerance, paced breathing, structured grounding — share the structural feature of reducing R_{req} in the moment.

X.2 Emotional tagging as retention-weight prior

The emotional valence $E(b)$ that biases Pass III sampling also serves as a retention-weight prior on Pass I pruning ([opt-theory.md](#) §3.6.5, closing paragraph). Patterns with high $|E|$ are flagged as relevance-relevant: they are over-sampled in adversarial testing and under-pruned in consolidation. This is the framework’s account of emotional memory enhancement, and it has direct clinical relevance. Emotionally salient content is *structurally* harder to revise; therapies that target the emotional tagging itself (re-evaluation, exposure, EMDR) work by reducing $|E|$ on the targeted content, which then makes ordinary maintenance able to do its job.

X.3 Flow states as optimal codec operation

Flow is the framework’s prediction for optimal operating conditions: R_{req} approaches C_{max} with minimal distortion, all bandwidth is in productive use, the

self-model recedes (its complexity is not needed at the moment), and the experiential signature is engagement without effort. Flow is therefore not “no thought” but *no spare capacity for self-monitoring*, which removes the introspective signal that effort would normally produce. The framework predicts that flow states should consolidate exceptionally well (Pass II runs cleanly afterward on patterns acquired at high R_{req}) and should be associated with reduced rumination during and after, since Pass III has no spare bandwidth during flow and may have less unresolved high- $|E|$ content to sample after.

XI. Empirical Predictions and Research Directions

XI.1 Falsification-style predictions

The framework supports a set of structurally specific predictions beyond those already pre-registered in [opt-theory.md](#) §6.8. These are stated here as *proposed predictions*, not formally pre-registered commitments — a future preregistration pass parallel to the core paper’s §6.8 will give them measurement specifications, falsification thresholds, and shutdown semantics. The intent is to make the document a research-programme prospectus, not a closed exposition.

Table 5: Proposed falsification-style predictions. Status pending formal pre-registration.

OPT psychology prediction	Possible measure	What would weaken it
Mind-wandering content is importance-weighted (surprise + threat), not base-rate-frequency-weighted, in both REM and waking low-load states	Experience sampling stratified by load + content coding for surprise and threat relative to base rates	Wandering content tracks base-rate frequency of recent activities and is not biased toward high- $ E $ branches
Productive mind wandering reduces future branch surprise or threat for the rehearsed branches	Experience sampling + later affective and surprise ratings on the same branches	Wandering content has no relation to later reduction in prediction error or threat appraisal
Rumination reflects elevated and uncalibrated β without compression gain	Repeated thought-content entropy + physiological arousal + unchanged belief-confidence across episodes	Rumination episodes reliably produce compression / generalisation gains comparable to productive reflection

OPT psychology prediction	Possible measure	What would weaken it
Sleep improves structured generalisation more than rote repetition	Rule-transfer tasks after sleep vs. equivalent wake, controlled for fatigue	Equal or greater benefit for unstructured repetition vs. structural generalisation after controlling fatigue
Low-load daytime windows improve nightly maintenance	Randomised “walk / shower / rest” windows vs. saturated schedules, with sleep architecture and next-day affect / cognition as outcomes	No effect on insight, affect regulation, sleep architecture, or intrusive thought reduction
Afternoon-timed maintenance practice outperforms pre-bed practice for sleep-disrupted populations	RCT of identical autogenic protocol at two times of day, stratified by baseline sleep quality	No timing effect, or pre-bed practice is at least as effective for sleep-disrupted populations
Externalised structured self-monitoring adds clinical effect beyond equivalent in-session time	RCT with and without structured note-taking, controlling for total intervention time and contact	No incremental effect from the note-taking arm
Therapy effects track OPT failure mode better than DSM category	Failure-mode stratification (consolidation-failure / precision-dysregulation / etc.) vs. diagnosis-only prediction of treatment outcome	DSM category predicts outcomes equally well or better than OPT failure-mode classification
Sleep restoration produces broad psychiatric improvement before symptom-targeted treatment fully takes effect	Cross-diagnostic trial with sleep-architecture-targeted intervention as primary	Sleep restoration has no precedence effect over symptom-targeted treatment in the conditions studied

XI.2 The Δ_{self} measurement problem

The framework predicts a structural limit on introspective self-report: any self-report is generated by the self-model, which is less complex than the system it models. This is itself a methodological constraint, not a defeat: it argues for the routine triangulation of self-report with behavioural, physiological, and external-observer data, and against research designs that depend on self-report alone for

assessment of internal state. The bandwidth-residual memo's $\Delta_{\text{self}}^{\text{op}} = \Delta_{\text{floor}} + \Delta_{\text{load}}$ decomposition suggests that the load-pressure term is the probe-able one: clinical research that systematically varies cognitive load while collecting both internal and external measures should reveal the size of the introspective gap.

XI.3 Operationalising compression gain

Several predictions in §XI.1 and the rumination-vs-reflection study sketched in Appendix A depend on a single load-bearing construct: *compression gain*. The framework currently uses the construct in a structural sense — a reduction in the Kolmogorov complexity of K_θ required to predict the same observation stream within tolerable distortion (formally, $\Delta K_{\text{compress}}$ of [opt-theory.md](#) Eq. T9-8). For empirical work this needs a proxy that can be measured on human-experiment time scales without direct access to $K(K_\theta)$.

A reasonable starting-point proxy combines three observables, required to move together over the same content: (a) **same or better task prediction** — lower error on the relevant prediction task before and after the candidate consolidation episode; (b) **lower subjective load** — reduced self-reported effort or fatigue on parallel content; (c) **lower physiological arousal during continued operation** — measured via heart-rate variability, electrodermal response, or cortisol depending on time scale. Compression gain is operationally indicated when all three move in the predicted direction. Partial movement counts as ambiguous and explicitly does *not* falsify the predictions of §XI.1 either way.

This is a sketch of one candidate proxy, not a validated measure. Formal validation, alternative proxies (e.g., entropy of subsequent thought content; transferability to novel instances of the same structural rule), and methodological decisions about which proxy to commit to for which §XI.1 prediction are deferred — catalogued in Appendix B.10 as future work, and to be settled in specific preregistration documents when those are written. The minimum methodological commitment in this paper is to flag the construct's load-bearing status explicitly, so that no §XI.1 prediction is read as testable without first settling the compression-gain operationalisation for that test.

XI.4 Clinical and self-experimental implications

The framework's clinical implications, where stated, are conservative in detail and ambitious in framing. In detail: the framework is compatible with approaches the evidence already supports — sleep restoration, structured low-load windows, exposure-based and reconsolidation-based treatments, structurally appropriate pharmacology, externalised self-monitoring. In framing: the orienting goal is *codec stewardship*, with symptom relief as a marker of progress rather than the sole target. The self-experimental implication is correspondingly conservative: deliberate attention to sleep, to maintenance windows, to the difference between productive and pathological forward-fan activity, and to the use of externalised supervisory signal (journals, trusted others, structured logs) where the self-model is provably insufficient.

XI.5 Limits of the present document

The treatment in this paper is deliberately intra-psychic. It therefore cannot, on its own, explain attachment injury, loneliness, social defeat, identity formation, cultural meaning, moral development, family systems, institutional trauma, or collective cognition. In the present document, these enter only as *inputs* to the individual codec, not as coupled-codec or system-level dynamics. This is a significant limit because depression, PTSD, addiction, dissociation, and psychosis are often deeply social in onset, maintenance, and recovery: any clinically complete account must extend the framework with apparatus that this paper does not develop. The basic machinery for that extension — inter-observer coupling under compression parsimony — is introduced in the core theory’s inter-observer-coupling apparatus, especially [opt-theory.md](#) Appendix T-10, but its translation into social, developmental, and cultural psychology is left to a separate companion. The present document is the intra-psychic foundation; what follows from coupling between codecs is a different and substantial undertaking.

A second limit is that several of the structural mappings in §VII reorganise diagnostic categories in ways that have not been clinically tested. The framework therefore offers a research programme, not a clinical taxonomy. Until the predictions in §XI.1 are tested, the OPT failure-mode framing should be held alongside, not above, existing diagnostic classifications.

XII. Conclusion: Toward a Maintenance-Oriented Psychology

A psychology read through OPT places the Maintenance Cycle at the centre. The framework does not replace the cognitive-science, clinical, or neuroscience literatures that supply the substance of what it describes; it offers a structural backbone under which much of what those literatures already establish — attention, memory, affect, motivation, sleep, and the disorders that arise when these go wrong — can be read as instances of what a finite codec must do to maintain a coherent stream against substrate noise on a finite budget. On that reading, clinical disorders are interpreted as failure modes of identifiable apparatus; therapeutic practices that work are those that support the apparatus rather than only labelling its symptoms; empirical predictions can be stated in falsification-ready form (§XI); and the structural limits on self-knowledge become precise.

The paper’s claim to relevance is modest. Much of the empirical substance is drawn from established or active scientific literatures. OPT contributes a vocabulary, a small set of structural commitments (§I.3), and a research programme. Whether that contribution earns its keep is an empirical question that the predictions in §XI exist to answer.

The treatment here is deliberately incomplete in one direction. Social, cultural, and interpersonal psychology — the coupling between codecs and the structures they build together — are not in scope, and the framework’s machinery does not

yet straightforwardly extend to them. That extension is a separate work. The intra-psychic account is the foundation it would rest on.

The framing also makes one ethical point worth restating. *Codec stewardship* — protecting the conditions under which the maintenance cycle can run — is a positive object of care, not merely the absence of pathology. The ethics paper develops this civilisationally; this paper develops it for the individual mind. Both readings sit alongside, not against, the precautionary frame: avoiding suffering remains necessary, and supporting the system that can avoid it remains the upstream condition.

References

References are local to this paper (self-contained numbering; the earlier convention of continuing `opt-theory.md`'s numbering was retired when the core list grew into the claimed range). The OPT framework primitives — the Maintenance Cycle apparatus, Δ_{self} (Conjecture P-4), Narrative Drift — are cited by section in `opt-theory.md` rather than by bracket number:

- [1] Revonsuo, A. (2000). The reinterpretation of dreams: An evolutionary hypothesis of the function of dreaming. *Behavioral and Brain Sciences*, 23(6), 877–901.
- [2] Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science*, 330(6006), 932.
- [3] Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., & Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science*, 315(5810), 393–395.
- [4] Andrews-Hanna, J. R. (2012). The brain's default network and its adaptive role in internal mentation. *The Neuroscientist*, 18(3), 251–270.
- [5] Buckner, R. L., Andrews-Hanna, J. R., & Schacter, D. L. (2008). The brain's default network: anatomy, function, and relevance to disease. *Annals of the New York Academy of Sciences*, 1124(1), 1–38.
- [6] Schultz, J. H. (1932). *Das autogene Training: Konzentrative Selbstentspannung*. Thieme.
- [7] Stetter, F., & Kupper, S. (2002). Autogenic training: a meta-analysis of clinical outcome studies. *Applied Psychophysiology and Biofeedback*, 27(1), 45–98.
- [8] Diekelmann, S., & Born, J. (2010). The memory function of sleep. *Nature Reviews Neuroscience*, 11(2), 114–126.
- [9] Walker, M. P. (2017). *Why We Sleep*. Scribner.
- [10] Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: a meta-analytic review. *Journal of Consulting and Clinical Psychology*, 78(2), 169–183.
- [11] Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a transdiagnostic process. *International Journal of Cognitive Therapy*, 1(3), 192–205.

- [12] Spinhoven, P., Klein, N., Kennis, M., Cramer, A. O. J., Siegle, G., Cuijpers, P., ... Bockting, C. L. H. (2018). The effects of cognitive-behavior therapy for depression on repetitive negative thinking: a meta-analysis. *Behaviour Research and Therapy*, 106, 71–85.
- [13] Foa, E. B., Hembree, E. A., & Rothbaum, B. O. (2007). *Prolonged Exposure Therapy for PTSD: Emotional Processing of Traumatic Experiences*. Oxford University Press.
- [14] American Psychological Association. (2017). *Clinical Practice Guideline for the Treatment of Posttraumatic Stress Disorder (PTSD) in Adults*.
- [15] Edinger, J. D., & Carney, C. E. (2014). *Overcoming Insomnia: A Cognitive-Behavioral Therapy Approach*. Oxford University Press.
- [16] Buzsáki, G. (2015). Hippocampal sharp wave-ripple: a cognitive biomarker for episodic memory and planning. *Hippocampus*, 25(10), 1073–1188.
- [17] Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., ... Corlett, P. R. (2018). The predictive coding account of psychosis. *Biological Psychiatry*, 84(9), 634–643.
- [18] Schwartenbeck, P., & Friston, K. (2016). Computational phenotyping in psychiatry: a worked example. *eNeuro*, 3(4), ENEURO.0049-16.2016.
- [19] Huys, Q. J. M., Maia, T. V., & Frank, M. J. (2016). Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature Neuroscience*, 19(3), 404–413.
- [20] Friston, K. J., Stephan, K. E., Montague, R., & Dolan, R. J. (2014). Computational psychiatry: the brain as a phantastic organ. *The Lancet Psychiatry*, 1(2), 148–158.
- [21] Domhoff, G. W. (2018). *The Emergence of Dreaming: Mind-Wandering, Embodied Simulation, and the Default Network*. Oxford University Press.
- [22] Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science*, 3(5), 400–424.
- [23] Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167(7), 748–751.
- [24] Cusack, R., Ranzato, M., & Charvet, C. J. (2024). Helpless infants are learning a foundation model. *Trends in Cognitive Sciences*, 28(8), 726–738. DOI: 10.1016/j.tics.2024.05.001.
- [25] Gomez-Robles, A., Nicolaou, C., Smaers, J. B., & Sherwood, C. C. (2024). The evolution of human altriciality and brain development in comparative context. *Nature Ecology & Evolution*, 8, 133–146. DOI: 10.1038/s41559-023-02253-z.
- [26] Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, 10(1), 89–96. DOI: 10.1111/j.1467-7687.2007.00569.x.
- [27] Köster, M., Kayhan, E., Langeloh, M., & Hoehl, S. (2020). Making sense of the world: infant learning from a predictive processing perspective. *Perspectives on Psychological Science*, 15(3), 562–571. DOI: 10.1177/1745691619895071.

- [28] Paulus, M. P., Feinstein, J. S., & Khalsa, S. S. (2019). An active inference approach to interoceptive psychopathology. *Annual Review of Clinical Psychology*, 15, 97–122. DOI: 10.1146/annurev-clinpsy-050718-095617.
- [29] Hamburg, S., et al. (2024). Active inference for embodied neuromorphic agents. *Entropy*, 26(7), 582. DOI: 10.3390/e26070582.

Forward-fan / threat-simulation reference [1] (Revonsuo) and Free Energy / active-inference references are already in [opt-theory.md](#). Cross-references to existing OPT bibliography use the core paper’s numbering; reference numbers here begin at 114 to continue without collision.

Appendix A: Preregistration sketch — rumination vs. productive reflection

The full prediction table in §XI.1 is a research-programme prospectus, not a pre-registered protocol. The framework will earn its empirical keep, or fail to, through specific studies that operationalise individual predictions tightly enough to be weakening conditions, not just narrative comparisons. This appendix sketches what the most promising **first** preregistered study would look like, on the recommendation that it should target the narrowest and most central distinction the framework draws: between rumination (Pass III stuck at high β without compression gain) and productive reflection (Pass III that reduces future surprise or threat for the rehearsed branches). Other predictions in §XI.1 are valuable, but this one is closest to the framework’s load-bearing claim that the same operator runs in both states and that the difference is mechanically identifiable.

This is a *sketch*, not a finished preregistration. It names what would need to be specified for a registered report; the actual specifications belong to a separate methodological work.

Element	Sketch for the first preregistered study
Hypothesis (OPT-derived)	Rumination episodes are characterised by high importance-weighting (β) sampling that <i>does not</i> reduce subsequent prediction error or threat appraisal for the rehearsed branches, whereas productive reflection episodes show measurable reduction. The two are distinguishable by mechanism, not only by valence.

Element	Sketch for the first preregistered study
Target population	Adults with current self-reported repetitive negative thinking sufficient to interfere with daily function, recruited without DSM diagnosis as the primary entry criterion. Stratification by depression and anxiety symptom severity. Exclusion criteria following standard clinical-research safety practice.
Conditions / arms	Within-subject contrast across episodes: (a) episodes classified as ruminative by content + self-report + behavioural marker; (b) episodes classified as productive reflection by the same multi-modal classification. Cross-classifier agreement reported.
Primary outcome	Pre-vs-post change in (i) confidence-weighted prediction error for the rehearsed branch on a structured probe administered immediately after the episode and at +24h, and (ii) self-rated threat appraisal for the same branch over the same intervals.
Operationalisation of “compression gain”	Reduction in the entropy of subsequent thought content sampled on the same topic, combined with reduction in time-to-resolution on a structured problem-solving probe scored by blinded raters. The operational definition is part of the preregistration, not derived post-hoc.
Statistical threshold	Effect sizes and confidence intervals reported; primary threshold set per registered-report convention (one-tailed test where the framework’s prediction is directional, two-tailed otherwise; alpha 0.01 per preregistered primary contrast; sample size powered to detect a medium effect).

Element	Sketch for the first preregistered study
What would weaken OPT	If episodes classified as rumination (by content + arousal + repetition) reliably produce post-episode reduction in confidence-weighted prediction error and threat appraisal comparable to productive reflection — i.e., the mechanistic distinction OPT draws is absent — then the §XI.1 prediction on rumination fails and the central β account is undermined. Failure of any single measure is not sufficient; the framework loses one specific prediction, not the whole programme.
What would not, by itself, weaken OPT	A null effect on a single measure that is statistically underpowered, or that fails to match the operationalised definition of compression gain. The preregistration commits to which failures count.

The framework’s broader claim is that a comparable preregistration logic should be applied to each prediction in §XI.1 as it is approached empirically. The rumination-vs-reflection study is the right first step because it isolates the operator-level claim (β dysregulation as the mechanism of pathological wandering) from broader claims about disorder categories, treatment protocols, or pharmacology — none of which the framework derives.

Appendix B: Future Work and Deliberate Deferrals

The v0.3 reviews of this paper recommended a number of additions that would extend the treatment into a fuller account of the healthy waking organism. Some of those additions are present in v0.4 (the codec-ontogeny section §II.5 and the compression-gain operationalisation sketch §XI.3); the rest are catalogued here as *deliberately deferred*.

The catalogue serves a real architectural function. A reader who notices that affect-beyond-threat, the action loop, or executive-function architecture is absent should be able to confirm that the absence is by design rather than by oversight, and should be able to see what a future version or sibling paper would need to add. Entries are listed roughly in order of structural priority: items most likely to be brought into a future version or companion paper are first.

B.1 The Embodied Codec. Interoception, allostasis, body load, fatigue, pain, illness, hormonal state, circadian phase, exercise, breathing, gut state, sexual

arousal, temperature. A foundational decomposition of R_{req} as $R_{\text{exteroceptive}} + R_{\text{interoceptive}} + R_{\text{proprioceptive}} + R_{\text{homeostatic}} + R_{\text{social/contextual}}$ would strengthen the existing sections on anxiety, depression, addiction, chronic pain, dissociation, and suffering. The interoceptive predictive-processing literature ([28], Seth’s interoceptive-inference programme) is the natural anchor. Deferred because a foundational treatment requires more careful integration with that literature than this version can responsibly attempt.

B.2 The Waking Control Cycle. The daytime complement to \mathcal{M}_τ : state estimate \rightarrow importance weighting \rightarrow policy selection \rightarrow action \rightarrow prediction-error update. Action, affordances, motor prediction, goal hierarchy, habit, skill mastery, and the relation between policy compression and ordinary behaviour. The current paper has the offline (maintenance) loop in formal depth and the online (action) loop only in the background of the active-inference framing. A positive theory of behaviour, not only of maintenance and breakdown, would belong here. Deferred because it constitutes a chapter at minimum and requires careful native-OPT framing (branch selection, policy compression, prediction-error-driven action) rather than imported ecological-psychology vocabulary.

B.3 Affect beyond threat and surprise. The current treatment subsumes emotion into $E(b) = -\log P_{K_\theta}(b|z_t) + \alpha \cdot \text{threat}(b)$, which captures threat and importance-weighting cleanly but does not develop joy, curiosity, boredom, meaning, grief, anger, shame, or disgust as first-class control signatures. A useful direction is to read positive valence as expected compression gain or policy-space expansion, and negative valence as expected overload, blocked policy, or compression failure. The full taxonomy is its own undertaking. Deferred because it requires the Waking Control Cycle (B.2) to land first.

B.4 Memory-systems taxonomy. Working, episodic, semantic, procedural, prospective, emotional, and autobiographical memory as different codec layers, each with its own characteristic failure modes. The current paper’s memory treatment is mostly via Pass II consolidation; a layered account would distinguish PTSD (consolidation failure on emotional/autobiographical content) from semantic confusion, depressive overgeneral memory, dementia-related autobiographical erosion, procedural habit lock-in, and prospective-memory failures more cleanly than the current framing does. Deferred to a future version.

B.5 Executive function and metacognitive scaffolding architecture. Inhibition, task switching, planning, error monitoring, uncertainty monitoring, cognitive flexibility, attentional set, meta-awareness, and external scaffolding as the codec’s policy-control layer for B_{max} allocation. Would unify mindfulness, CBT, journaling, and structured routines as different forms of metacognitive scaffolding rather than isolated therapeutic exemplars. Deferred because it requires the Waking Control Cycle (B.2).

B.6 Individual differences as parameter-space variation. B_{max} , β , λ , precision priors, interoceptive gain, Pass III bias, maintenance efficiency, self-model rigidity, scaffolding dependence — read as personality parameters rather

than only as clinical-variant levers. Big Five-style mappings are easy to write and easy to over-claim; a parameter-space framing without committal mappings would let the framework speak to ordinary individual variation. Deferred because the empirical work to validate the parameter space is itself a project.

B.7 Normal-psychology positives. Curiosity, play, creativity, humour, flow, aesthetic experience, skill mastery, meaning, resilience, ordinary problem-solving, and insight as expressions of well-running apparatus. The current paper’s §X.3 (flow) and the codec-stewardship framing in §0.2 and §XII gesture toward this, but the area deserves its own treatment. Deferred because it requires affect (B.3) and the Waking Control Cycle (B.2) to land first.

B.8 Boundary states as natural stress tests. Anesthesia, delirium, mania, psychedelics, hypnosis, deep meditation absorption, panic attacks, chronic pain, depersonalisation, grief, burnout, and severe sleep deprivation each stress a different part of the OPT apparatus (anesthesia probes whether $P_\theta(t)$ disappears or becomes inaccessible; delirium tests high-noise low-coherence K_θ ; psychedelics test relaxed priors and altered precision; mania tests runaway policy-space expansion under reduced pruning; chronic pain tests interoceptive prediction lock-in; flow tests unusually efficient action-prediction coupling). A short catalogue mapping each state to which apparatus it probes would make the framework more empirical-feeling without committing to specific clinical mechanisms. Deferred because each state has its own contested literature.

B.9 Perceptual psychology. Perceptual learning, illusions, attentional and change blindness, body ownership illusions, affordance perception, active sensing, sensory substitution, phantom limb, the hallucination-imagery-perception continuum, and pain as perception. The current paper’s §II.2 has the predictive-construction story but a dedicated perceptual chapter would bridge ordinary perception, illusion, hallucination, and psychosis more gracefully. Deferred to a future version.

B.10 Full operationalisation methods appendix. §XI.3 sketches one proxy for compression gain. A full methods appendix would supply operational definitions of R_{req} , β , compression gain, Pass III bias, pruning, and consolidation; candidate behavioural / physiological / sleep / experience-sampling / clinical-scale measures for each; minimum viable preregistered studies (Appendix A is the first of these); and explicit thresholds for what counts as failure. This is the work that turns the prediction table of §XI.1 into a research programme proper.

B.11 The coupled-codec / social companion. Most of the deferred social, cultural, developmental, and interpersonal psychology requires inter-observer coupling apparatus introduced in [opt-theory.md](#) Appendix T-10. A separate companion paper would treat: interpersonal psychology, attachment, family systems, group dynamics, cultural psychology, developmental psychology beyond intra-codec ontogeny (which §II.5 covers as far as this paper goes), social identity, moral psychology beyond suffering, and educational, organisational, and political psychology. As an interface contract, the intra-psychic paper exports to that

future companion the following codec-state variables: K_θ stability, R_{req} baseline, β calibration, λ retention threshold, Pass III content bias, self-model rigidity, external scaffolding dependence, and $\Delta_{\text{self}}^{\text{op}} = \Delta_{\text{floor}} + \Delta_{\text{load}}$ under varying load. The future companion’s central question is then: what happens when two or more codecs regulate each other’s prediction error?

B.12 Virtual-standing-state compatibility note. Open item carried from the (now-archived) virtual-standing-state work, landed as core [opt-theory.md](#) §8.6.1: under the fully-virtual reading, $P_\theta(t)$ and \mathcal{M}_τ are structural properties a filter-passing stream *has*, not machinery it *runs*. This paper’s intra-psyctic treatment uses the operational reading throughout and is unaffected (the dual reading does not change any clinical mapping or magnitude). A short global neutrality sentence to that effect should be added to §0.4 / §III.1 in a follow-up pass; deferred here as low-priority housekeeping.

This list is not exhaustive. It marks the most prominent items raised during the v0.3 review process; future review may add or remove entries.

Version History

Table 6: Revision History.

Version	Date	Summary
0.1	May 23, 2026	Initial draft. Psychological translation of \mathcal{M}_τ (§§I–III); mind wandering and rumination as Pass III states (§§IV–V); neuroscience as substrate bridge (§VI); psychiatric failure-mode map across nine categories (§VII); therapeutic interventions as codec hygiene (§VIII); agency, suffering, flow (§§IX–X); empirical predictions and conclusion (§§XI–XII). Bundled with opt-theory.md ; shares core DOI. Intra-psyctic scope.

Version	Date	Summary
0.2	May 23, 2026	<p>OpenAI review absorbed. Abstract restructured into four labelled paragraphs (Purpose, Core mapping, Clinical mappings, Scope and posture) with explicit statement of purpose. Dedicated non-clinical / non-diagnostic blockquote disclaimer at the end of the abstract. Subtitle: “Intra-Psychic Psychology and Psychiatry”. Added §0.3 Claim Status Table.</p> <p>Humility pass: §I.3 substantially expanded to lead with explicit acknowledgement of source literatures; OPT’s distinctive contributions restated as a small numbered set (1)–(5); §VII opener acknowledges that categories, phenomenology, differential diagnoses, and treatment evidence come from clinical psychology and psychiatry, not from OPT; §XII conclusion softened to “the substance is established science.” Tone shift to “is modeled as” across abstract and §VII. $P_\theta(t)$ aligned with standing-state vs. update-channel distinction (§II.2). Δ_{self} hedged to Conjecture P-4 (§II.4, §VII.5). REM $R_{\text{req}} \approx 0$ corrected (§III.1, §VI.3). Dreaming hedged to “an important component” (§III.1, §VI.3); competing dream theories acknowledged. Psychosis rewritten in predictive coding register</p>

Version	Date	Summary
0.3	May 23, 2026	<p>Second OpenAI review absorbed.</p> <p>Epistemic-hygiene pass: §I.3 “most sentences are restatements” softened; §I.2 cross-reference corrected</p> <p>(opt-theory.md §3.8 → Appendix T-10 for inter-observer coupling); §VII opener “decades organised around predictive-coding” softened to “an increasing subset”; §VII.10 “structural diagnoses” → “structural characterisations / failure-mode hypotheses”; §VI.4 pharmacology caveat mirrors §VIII.4; §VII.5 dissociation Δ_{self} wording → “first-person continuity and agency”; §VII.7 addiction broadened from substance to “addictive reinforcer” covering behavioural addictions; §VII.8 ADHD “well-documented hyperfocus” → “commonly reported”; §VIII.2 afternoon-timing claim downgraded to “worth testing”; §VIII.5 renamed “Sleep restoration and CBT-I as maintenance supports” with scope softened; §XII conclusion: “substance is established science” → “much of the empirical substance is drawn from established or active scientific literatures”. Additions: §0.4 <i>How to read the mappings</i> callout; §0.5 plain-language glossary (Table 2) for psychology readers;</p>

Version	Date	Summary
0.4	May 23, 2026	<p>High-level review absorbed selectively, motivated by parallel work in <code>opt-ai-design.md</code> §7.4 (“Forced developmental curriculum”). New §II.5 <i>Codec ontogeny</i>: intra-psychic developmental story — sensory-motor bootstrap, core knowledge and object permanence, body schema formation, autobiographical memory emergence, adolescence as self-model refactoring, aging as \mathcal{M}_τ degradation, dementia and amnesia as model/residual dissociation. Bidirectional cross-references with <code>opt-ai-design.md</code> §7.4 (internal companion paper). New §XI.3 <i>Operationalising compression gain</i>: sketches one candidate proxy (task prediction + subjective load + physiological arousal) for the load-bearing construct that §XI.1 predictions and Appendix A depend on; existing §XI.3/§XI.4 renumbered to §XI.4/§XI.5. New Appendix B <i>Future Work and Deliberate Deferrals</i>: eleven-entry catalogue (Embodied Codec, Waking Control Cycle, affect beyond threat, memory-systems taxonomy, executive function, individual differences, normal-psychology positives, boundary states, perceptual</p>

Version	Date	Summary
0.5	May 23, 2026	<p>Third OpenAI review absorbed. New §II.6 <i>Artificial-consciousness bridge</i>: consolidates the AI-design implications previously scattered across §II.5, §VI.5, §VII into a compact table (Table 4) plus dependency diagram, with the load-bearing claim “a conscious-capable codec is not merely architected; it is developed and maintained” stated explicitly.</p> <p>Cross-references to <code>opt-ai-design.md</code> §§5.5/5.6/5.8/6.1/6.3/7.4/7.5/7.6/9.6.</p> <p>Wording softening: §II.5 altriciality necessity claim → “plausible biological solution” rather than “structural requirement”; §II.5 dementia/amnesia → “OPT-internal reading of ambiguous clinical phenomena, not a direct measurement of Δ_{self}”; §II.5 “born-mature” deployment claim explicitly relabelled as a design hypothesis, not an engineering rule; §VII.9 sleep-restoration scope tightened to “when sleep disruption is part of the maintaining loop”.</p>

Version	Date	Summary
0.6	May 23, 2026	<p>v0.5 verification housekeeping. Abstract: “sleep hygiene” → “sleep restoration”; “social, cultural, developmental, and interpersonal” scope sentence updated to read “social, cultural, interpersonal, and developmental psychology beyond intra-codec ontogeny”, reflecting §II.5. §I.2 developmental bullet now says “developmental psychology beyond the single-codec ontogeny sketched in §II.5”. §XI.5 cross-reference regression fixed: <code>opt-theory.md</code> §3.8 → Appendix T-10 (same correction that landed for §I.2 in v0.3 but had been re-introduced in §XI.5). Duplicate table numbers renumbered: §XI.1 falsification table 2 → Table 5, Revision History table 3 → Table 6 (collisions caused by Tables 1–4 in §0.3 / §0.5 / §I.3 / §II.6 being numbered after the originals). §XI.1 body and Table 5 caption: stale “v0.3 pass” / “pending formal pre-registration in v0.3” → “future preregistration pass” / “Status pending formal pre-registration”. Table 4 AI bridge: “software-only Markov blanket” → “merely a declarative self-model or transient context window” for clarity.</p>

Version	Date	Summary
0.7	June 2026	Core-v4.1.x alignment: §I.4 Δ_{self} reframed locus → capacity gap + individuation (no chooser in the gap); §I.1 gloss + fully-virtual idiom pointer (theory §1.6/§8.6.1).
0.8	June 2026	References made self-contained: the shared-numbering convention had crashed (local [114]–[141] collided with core refs [114]–[118] added in v4.1.x); renumbered to local [1]–[29], Revonsuo added as [1].
0.9	June 2026	Purpose statement aligned to capacity vocabulary: “structural self-model incompleteness” → “a budgeted self-model capacity gap (structural self-model incompleteness, Conjecture P-4)”.