No. 2026-007

When Memory Governs Memory: An Endogenous-Operator Model of Recursive Psychological Activity

August 18, 2026

Takanori Suzuki

When Memory Governs Memory: An Endogenous-Operator Model of Recursive Psychological Activity

Takanori Suzuki*

August 18, 2026

Research Institute of Business Administration (Sanken), Waseda University
Working Paper Series, No. 2026-007

Working paper note. This paper is an English version of the author's Japanese-language article, forthcoming in *The Waseda Commercial Review* (*Waseda Shogaku*), No. 473 (December 2026). It is circulated in the Working Paper Series of the Research Institute of Business Administration (Sanken), Waseda University, with the permission of the Waseda Commercial Studies Association, publisher of *The Waseda Commercial Review*, granted on August 18, 2026. A revised English version may be submitted to a journal.

Abstract

This paper develops a theoretical model in which psychological distress is recursively complicated not only by memory contents but also by operators that memory endogenously generates to conceal, modify, punish, justify, or control itself. Memory is defined broadly as the accumulated past that is currently active, including experience, knowledge, evaluation, self-images, expectations, and response tendencies. The paper distinguishes functional memory, which is used by a present task, from self-referential psychological memory, which uses the present to preserve or modify a self-image. An operator-generating map from memory states to memory-processing rules is introduced to formalize recursive layers such as punishment, punishment of the punisher, and monitoring of the higher-order controller. The controller is therefore not assumed to be an independent agent but is represented as an endogenous operator generated from the same memory state. Attention is modeled not as a superior controller but operationally as functional non-closure: memory-generated outputs are not re-entered into the system as present facts. Analytical results and numerical illustrations establish three points. First, in the non-closure limit memory is not erased: attention only suspends rule formation and the authority of rules over the present; the rule-forming history is preserved, registration into non-recursive records continues, and the fate of the active content is governed by the natural dynamics alone. Second, in a one-dimensional specialization, quieting and persistent self-amplification are

*Faculty of Commerce, Waseda University. Email: suzuki@waseda.jp. ORCID: 0009-0008-8955-7003.

separated in the critical regime by an explicit threshold of accumulated history, so that identical parameters can yield qualitatively different outcomes depending on history. Third, for the general model, an explicitly checkable condition on the primitives guarantees convergence of the recursive hierarchy. The paper discusses implications for psychopathology, treatment-induced control loops, and empirical testing, and offers a preliminary, testable framework free of religious or metaphysical premises.

Keywords: memory; self-reference; endogenous operator; rumination; thought suppression; attention; nonlinear dynamics; formal theory; psychiatry
JEL classification: C61, D91, I12.

1 Introduction

People become entangled in mental activities such as anxiety, tension, rumination, self-evaluation, achievement striving, defensiveness, shame, and concealment. Yet distress is not produced by primary mental activity alone. A secondary movement arises that tries to "do something about" the primary activity, and this control effort generates new evaluations, feelings of failure, monitoring, and punishment that can strengthen the very activity it targets. When bodily tension arises before a class, for example, the thought "the students must not notice" may follow, and one begins to control one's face, voice, and posture. The control stiffens the body further; the stiffness is re-entered as evidence that "they will surely notice"; the concealment effort escalates. Here the movement that tries to solve the distress has entered the state equation of the distress itself.

In a companion paper published in Japanese, mental activity was modeled as a memory-based self-amplifying loop whose effective gain is reduced by self-observation, whereas the intention to "quiet the mind" forms a further positive feedback of its own (Suzuki, 2026). That paper, however, left unspecified where the "trying-to-quiet" state comes from, and how such a control state generates higher-order controllers, punishers, and justifiers. It also posed, as its next task, the question of how entry into pure self-observation might be assisted from outside (an externally assisted transition model). Before the conditions of such assistance can be designed, one must first clarify how the state to be assisted—the trying-to-quiet state—is generated and updated from within the memory system. The present paper is positioned as the intermediate theory connecting the companion paper's comparative model with the future externally assisted transition model.

The point of departure is the refusal to posit the controller as an agent independent of what it controls. In the psychological domain, "the I that tries to erase anxiety," "the I that punishes the self that conceals," and "the I that judges the punishing self as violent" may all be memory movements formed from the same past, self-images, evaluations, fears, and norms. Memory then does not merely store contents; it generates the rules—the operators—that process those contents. Moreover, the results of each operation are accumulated into the next memory state and rewrite the next operator itself. The defendant, the prosecutor, the judge, the internal auditor of the judge, and even the reviser of the penal code all emerge endogenously from the same system.

The purpose of this paper is to formalize this structure minimally as an *endogenous-operator model*. The contributions are fourfold. First, functional memory and self-referential psychological memory are distinguished not by content but by the direction of servitude between memory and the present. Second, a map that generates concealment, punishment, justification, and control operators from memory states is introduced, and the recursive escalation is represented as a history-dependent nonlinear system. Third, attention is represented not as a topmost controller but as *functional non-closure*: the condition that memory-generated outputs are not re-entered into the system as present facts. Fourth, implications are derived for case formulation, treatment-induced aggravation, self-blame, and within-person time-series testing in psychiatry and clinical psychology. All formal results require only elementary analysis, and full proofs are given in the text and the appendix.

This paper does not personify memory. Expressions such as "memory conceals" or "memory punishes" do not attribute intention or malice to memory. They are a convenient description of a structure in which the repetition of local self-preservation, prediction, explanation, and control comes to look as if a cunning agent were at work. Nor does the paper offer a comprehensive etiology or therapy for mental disorders; it is a theoretical hypothesis describing one facet observable in everyday and clinical life.

2 Related work and positioning

2.1 Rumination, unresolved goals, and thought suppression

Research on rumination has shown that negative thinking repeats, persists, and is maintained by self-referential processing (Smith and Alloy, 2009; Nolen-Hoeksema et al., 2008; Watkins, 2008). Cueing an unresolved personal goal causes ruminative self-focus to persist, consistent with the control-theoretic view that the discrepancy between the present and a goal state triggers repetitive thought (Roberts et al., 2013). Two strands of this literature anticipate specific components of the present model. The habit–goal framework construes trait rumination as a mental habit and links response-styles theory to control theory through the relation between habits and goals (Watkins and Nolen-Hoeksema, 2014); the automaticity and history dependence of the habitual rule correspond, in the present model, to the operator gain that grows with the rule-forming history. The impaired-disengagement hypothesis locates pathological rumination in a failure of attentional disengagement from self-referential content (Koster et al., 2011), an attentional counterpart of what is here called closure. The perseverative-cognition hypothesis extends the health costs of such repetitive activation beyond mood (Brosschot et al., 2006).

Attempts to suppress unwanted thoughts can increase their recurrence and accessibility, as shown in classic experiments and meta-analysis (Wegner et al., 1987; Abramowitz et al., 2001). The finding demonstrates that "trying to control" does not always diminish the target and that the control effort itself can become a maintaining factor. Because suppression effects vary with task, target, measurement, and individual differences, this paper does not treat all control as harmful. What matters is self-referential control: the repeated correction of present inner experience against a self-image, a fear, or an ideal state.

2.2 Metacognitive models and decentering

The metacognitive model of Wells and Matthews (1994) and Wells (2009) identifies a cognitive–attentional syndrome—worry, rumination, threat monitoring, and unhelpful coping strategies—as the maintaining engine of emotional disorder. This lineage is closest to the present paper in making clear that the problem lies not only in thought contents but in the mode of responding to and controlling thoughts. In the S-REF model, however, coping strategies and procedural beliefs are stored in long-term memory and are selected and driven from there, so the mere memory origin of control rules is not what is new here. The differences are the following. First, control operators are generated state-dependently from the current history state (the explicit generating map Γ). Second, the results of each operation are re-entered into the history so that the next operator itself is updated (self-modification of the rules). Third, the controller is not an independent executive but an endogenous operator generated from the same memory state. Fourth, attention is treated not as another control strategy but as functional non-closure. Fifth, quantitative consequences—thresholds, history dependence, and basins of attraction—are derived from this structure (Section 6).

Decentering research, meanwhile, has organized meta-awareness of inner experience, disidentification from experience, and reduced reactivity to thought content as distinguishable processes (Bernstein et al., 2015, 2019). This connects with the present notion of attention, in which thought contents are not erased but noticed *as* thoughts, whereby their reactivity declines. Mind-wandering research likewise shows that consciousness is not fixed on a single object but is a dynamic process of excursion and re-noticing (Smallwood and Schooler, 2015).

2.3 Cognitive fusion, control-theoretic accounts, and Morita therapy

The clinical construct closest to the present notion of non-closure is cognitive fusion/defusion in Acceptance and Commitment Therapy (Hayes et al., 1999). Fusion is the mode in which the output of thought is experienced as present fact; defusion is the mode in which it is experienced *as* thought. The pair corresponds partially to high versus low values of the closure coefficient c in this paper (the fusion construct stresses the coupling of thought and overt behavior, so the correspondence with "re-entry as present fact" is not exact). For operationalization, the Cognitive Fusion Questionnaire (Gillanders et al., 2014) is a candidate; being a trait measure, it is complemented for within-day dynamics by the State Cognitive Fusion Questionnaire (Bolderston et al., 2019). The contribution here is to formalize this construct as a dynamics of operator generation and re-entry and to derive conditions for convergence and divergence (Sections 5 and 6).

From the control-theoretic side, Perceptual Control Theory conceives behavior as the hierarchical control of perception (Powers, 1973), and feedback-loop accounts of self-regulation brought the discrepancy-reducing loop into personality, social, and clinical psychology (Carver and Scheier, 1982). Higginson et al. (2011) used Perceptual Control Theory to give an integrative mechanistic account of psychological distress and therapeutic change, in which distress arises from conflict between control systems and change involves redirecting awareness to a level above the problem, with recovery driven by (partly random) reorganization. The present model differs on precisely the two points at issue: in Perceptual Control Theory the architecture of controllers is fixed and conflict between given controllers does the pathological work, whereas here the operators themselves are generated and rewritten from an accumulating history; and awareness redirected "upward" is still an operation within the hierarchy, whereas non-closure is not a level but the severing of re-entry. The finding that a changed relation to thought content (metacognitive awareness), rather than content removal, prevents depressive relapse (Teasdale et al., 2002) is consistent with the therapeutic implication of non-closure, which does not require erasure of content.

In the Japanese context, the correspondence with Morita therapy is suggestive. Morita's psychic interaction (*seishin kōgo sayō*), "contriving" (*hakarai*), and "as-it-is-ness" (*arugamama*) can plausibly be redescribed, in the present vocabulary, as the positive-feedback loop, the control-system operators, and non-closure, respectively (Morita, 1928/1998). Processes in which an emotional reaction becomes the object of a further emotional reaction—fear of fear (Goldstein and Chambless, 1978) and depression about depression (Teasdale, 1983)—have been described repeatedly in the clinical literature and provide phenomenological instances of the "punishment of punishment" treated in Sections 3 and 4. Intervention research on self-criticism and shame (Gilbert and Procter, 2006) targets this recursive layer in practice.

2.4 Symptom networks, dynamical systems, and formal theories

Network research on psychopathology understands disorders not as effects of a single latent cause but as systems of interacting symptoms with feedback—worry breeds insomnia, insomnia breeds fatigue, fatigue breeds further worry (Borsboom and Cramer, 2013; Borsboom, 2017). The present paper shares this dynamic viewpoint in treating distress, control effort, bodily constraint, and self-evaluation as interacting states. Two further developments in this tradition frame the present contribution. First, the emerging program of *formal theories* of psychopathology replaces verbal theories with explicit computational models; its flagship example formalizes a network theory of panic disorder as a nonlinear dynamical system and evaluates what the theory can and cannot explain (Robinaugh et al., 2024), and its methodology has been articulated as a general research strategy (Haslbeck et al., 2022). The present paper is an exercise in exactly this genre: a verbal clinical intuition—that the controller is generated by what it controls—is converted into equations whose consequences (thresholds, incubation, path dependence) can be checked.

Second, dynamical-systems work on mood has argued that the mood system can possess alternative stable states separated by tipping points, with critical slowing down—elevated autocorrelation and variance—as an early-warning signal of transitions into and out of depression (van de Leemput et al., 2014); elevated emotional inertia likewise predicts maladjustment (Kuppens et al., 2010). The threshold and basin results of Section 6 may look superficially similar, but their origin differs: bistability is not assumed at the level of mood states; instead, an explicit critical value of *accumulated history* is derived from the pathway "history $\rightarrow$ operator gain," together with theorems on incubation length and initial-value dependence. The two perspectives are complementary rather than competing, and Section 8 shows that the present model in fact predicts critical-slowing-down signatures as the effective gain approaches one.

2.5 Attention as precision control: hierarchical predictive processing

A different formal tradition models attention within hierarchical predictive processing and active inference. In the deep parametric active-inference model of Sandved-Smith et al. (2021), mental action is policy selection over higher-level cognitive states: attentional states modulate the precision of the mapping at the level below, and meta-awareness states modulate precision one level above that. This is the most explicit current formalization of attention *as* higher-order control, and it is exactly the architecture whose regress problem is analyzed in Remark 5.5 below: if the attention operator is itself generated from system states, a further level can always be asked for. Relatedly, Laukkonen and Slagter (2021) give a unifying predictive-processing account of meditation as the progressive reduction of counterfactual, temporally deep processing, mechanistically the closest neighbor of non-closure in the literature; there, however, the change is implemented as a reallocation of precision (gain), whereas here it is the severing of the re-entry of operator outputs as present facts, and the threshold and history theorems of Sections 5–6 have no counterpart in that account. On the phenomenological side, the contemplative-science construct of dereification—experiencing thoughts as processes rather than as direct reflections of reality (Lutz et al., 2015)—provides, alongside decentering and defusion, a further construct-validity anchor for the closure coefficient.

2.6 The remaining problem

The literature thus supplies important parts: repetitive thought, the paradox of suppression, metacognitive control, decentering, symptom networks, formal theories, and precision-based attention. What this paper focuses on is the point that the *control rules themselves are generated from the same memory state as the controlled content.* Not only memory contents but the operators of concealment, punishment, justification, and monitoring are endogenized, and the results of their operation update even the next operator. Furthermore, if attention is introduced as another higher-order controller, an infinite regress arises: was that controller not also generated by memory? To avoid this, attention is represented here not as an operator but as a condition on the functional closure of the loop.

The literature review above remains a targeted one for a theory paper; it does not claim complete novelty for the endogenous-operator construct, and a fully systematic review of adjacent constructs remains future work.

3 Basic concepts

3.1 An operational definition of memory

Memory, in this paper, is not mere stored recollection but the totality of the past that is acting on present perception and response: experience, knowledge, evaluation, self-images, expectations, fears, comparisons, norms, and response tendencies. In this sense memory is not a repository but a state operating in the present.

Decompose the memory state as

$$M_t = (M_t^{\mathrm{F}}, M_t^{\mathrm{S}}), \tag{1}$$

where M_t^{F} is functional memory and M_t^{S} is self-referential psychological memory. The two are not fixed partitions of content: the same remembered content can operate functionally in one situation and psychologically in another.

Definition 3.1 (Functional memory)**.** Functional memory is memory that the present task or environment uses as needed for responding. It may contain reference to the self, provided the reference is local and indexical and does not serve the maintenance or defense of a self-image.

For instance, "I tried this transformation yesterday" or "last week I lectured up to Chapter 3" contain self-reference, but as long as the present mathematics or teaching is using memory as a tool, they are functional.

Definition 3.2 (Self-referential psychological memory)**.** Self-referential psychological memory forms a self-image from the past and evaluates, filters, and controls the present against that self-image. Here it is not the present that uses memory; memory, through the "I," uses the present for self-preservation, self-expansion, and self-correction.

For instance, "the students must not notice that I am tense" and "an I that cannot do this proof lacks ability" evaluate present facts in relation to a self-image and trigger control.

3.2 Three kinds of self-reference

Self-reference can be distinguished into at least three layers.

(1) *Indexical self-reference*: the "I" that marks time, history, or bodily state. Example: "I did this calculation yesterday."

(2) *Evaluative self-reference*: comparing the present self with a self-image or norm. Example: "the I that failed is worthless."

(3) *Recursive self-reference*: further correcting, punishing, or monitoring the evaluative self-reference. Example: "the self that blames itself is violent, so that too must be stopped."

This paper is mainly concerned with the third layer. In recursive self-reference, the processing introduced to solve a problem becomes a new problem state, and each act of processing can raise the order of the system.

3.3 Endogenizing the operators

Let the memory space $\mathscr{M}$ be a Banach space equipped with a closed convex cone $\mathscr{M}_+$, and let $\mathcal{B}(\mathscr{M})$ denote the bounded linear operators on it. The map generating operators from memory states is

$$\Gamma : \mathscr{M} \longrightarrow \mathcal{B}(\mathscr{M}). \tag{2}$$

$\Gamma(M)$ represents the processing rules—concealment, punishment, justification, control, monitoring—activated by the state M. Typically,

$$\Gamma(M) = w_H(M)\mathsf{H} + w_P(M)\mathsf{P} + w_J(M)\mathsf{J} + w_C(M)\mathsf{C}, \tag{3}$$

where H, P, J, and C are basis operators of concealment, punishment, justification, and control, and the weights $w_k(M)$ depend on the memory state. The essential point is that the operators are not fixed exogenously but are generated and selected by the memory state itself.

3.4 Attention and functional non-closure

Attention is not defined here as a new superior operator. If attention is introduced as an operator $\mathcal{A}$ and one writes "$\mathcal{A}$ suppresses P," the question arises where $\mathcal{A}$ came from. If $\mathcal{A}$ too is generated from memory, it is merely another controller belonging to the same hierarchy.

Instead, let $c_n \in [0,1]$ (in continuous time, c_t) denote the degree of functional closure at each stage. $c_n = 1$ represents complete closure, in which operator outputs are re-entered essentially as they are; $c_n = 0$ represents complete functional non-closure, in which memory may well appear but its output is not re-entered as the present itself. For empirical work, closure is operationalized through an observable index A that covaries with it (measures of decentering and nonreactive awareness; Section 8):

$$c : [0,1] \longrightarrow [0,1], \qquad c'(A) < 0, \qquad c(0) = 1, \quad c(1) = 0, \tag{4}$$

with c continuous.

The order of definition deserves emphasis. The primitive concept of this paper is the degree of functional closure c_n itself; the first-person experience corresponding to low c_n is what this paper calls attention. A_t is an operational variable bundling observable indices that covary with c_t (decentering, nonreactive awareness; Section 8); the generative mechanism of attention itself is outside the explanatory scope of this paper and is explicitly left to the externally assisted transition model (Section 1) as the task of the next paper.

This definition does not mean that attention erases memory. Even when the memory "the students must not notice my tension" has appeared, there is a mode in which, being seen as a movement of memory, that memory does not become the sovereign that selects the next response; the definition captures this mode operationally. The statement has a literal formal counterpart in Section 5 (Proposition 5.1): in the non-closure limit $c_n \equiv 0$, only rule formation and the action of rules on the present stop; the rule-forming history is preserved, registration into non-recursive records continues, and whether the active content subsequently decays or persists is left to the nature of the natural dynamics. "Seeing" is thus not a strong form of suppression but an operation of a different kind from suppression (direct reduction of content).

4 The endogenous-operator model

4.1 Discrete-time model

The temporal evolution of psychological memory requires distinguishing channels with different roles. The first is the **natural dynamics**: the process by which active content changes and decays without operator intervention. The second is the **action channel**: rules generated from history act on present content. The third is the **rule-forming channel**: content is registered into a self-referential history and becomes material for future rules. In addition, there is a pathway by which content is stored as a non-recursive **record** (the storage of functional memory). What attention (non-closure) severs is the second and the third—the closed circuit that receives outputs as present facts and re-accumulates them into rule formation—not the natural dynamics and not the non-recursive record.

To make the distinction explicit, the state is divided into three. The **activation** at layer n, $M^{(n)} \in \mathscr{M}$ (the currently operating content; it includes primary perceptual and bodily states and is the object processed by self-referential operators; layers $n \geq 1$ are mostly self-referential). The **rule-forming history** $\mathcal{H}_n \in \mathscr{M}_+$. And the **non-recursive record** $\mathcal{K}_n \in \mathscr{K}$ (stored memory corresponding to the storage of functional memory; $\mathscr{K}$ is a space equipped with addition whose internal structure is not further specified; "non-recursive" refers to the role of the record in this model—records are not fed back into the operator-generating recursion—and implies neither that recorded contents are non-self-referential nor that records are accurate, neutral copies).

Let $D \in \mathcal{B}(\mathscr{M})$ represent the natural dynamics and set $\delta := \|D\|$ (in the numerical examples $\delta < 1$: natural decay; the role of this assumption is made explicit in Section 5). Let $\Psi : \mathscr{M} \to \mathscr{M}_+$ be a registration map into the history satisfying $\|\Psi(x)\| \leq \|x\|$ (nonlinearity is allowed), and let $\Theta : \mathscr{M} \to \mathscr{K}$ be the registration map into the record. The operator active at stage n is

$$\mathcal{O}_n = \Gamma(\mathcal{H}_n), \tag{5}$$

and the state is updated by

$$M^{(n+1)} = D[M^{(n)}] + c_n\, \mathcal{O}_n[M^{(n)}], \tag{6}$$

$$\mathcal{H}_{n+1} = \mathcal{H}_n + c_n\, \Psi(M^{(n+1)}), \qquad \mathcal{H}_0 = \Psi(M^{(0)}), \tag{7}$$

$$\mathcal{K}_{n+1} = \mathcal{K}_n + \Theta(M^{(n+1)}). \tag{8}$$

The closure coefficient c_n multiplies, in (6), the channel through which rules act on the present, and, in (7), the channel through which content is registered into the rule-forming history. Registration into the record channel (8) is not subject to closure: the registration of events continues. Since $\mathcal{H}_{n+1} - \mathcal{H}_n = c_n\Psi(M^{(n+1)}) \in \mathscr{M}_+$, the history accumulates monotonically in the cone order; this is what "cumulative" means in this paper. Ψ may be interpreted as registering the intensity and salience of episodes into the history rather than the signed content itself (accumulation without cancellation).

A terminological note. "Memory" in this paper refers, by context, to (i) the activation $M^{(n)}$, (ii) the rule-forming history $\mathcal{H}_n$, or (iii) the record $\mathcal{K}_n$; where confusion could arise, the term is qualified. The later statement "memory is not erased" (Proposition 5.1) concerns (ii) and (iii): under complete non-closure $c_n = 0$, what stops is registration into (ii) and the action of rules; registration into (iii) continues; and the fate of (i) is left to the natural dynamics D. Since $\mathcal{K}$ does not recurse into the self-referential loop that is the subject of this paper and corresponds to the storage of functional memory (Section 3), it is not treated explicitly in the analysis below. In reality there can be a pathway by which recorded contents are later recalled and used self-referentially and flow into the history ($\mathcal{K} \to \mathcal{H}$); this coupling is omitted here (Section 10).

Multiplying the history registration (7) by c_n expresses the theoretical position that "under pure observation, contents are registered as events (into $\mathcal{K}$) but are not accumulated into self-referential rule formation ($\mathcal{H}$ does not grow)." If one adopts instead the position that factual records also contribute to rule formation, the variant $\mathcal{H}_{n+1} = \mathcal{H}_n + \Psi(M^{(n+1)})$ may be used. Under this variant, Proposition 5.2, Corollary 5.3, and Proposition 6.1, which use only norm estimates on the M-side, remain valid as stated, and the dichotomy of Section 6.3 (Theorem 6.2) also holds qualitatively as the case $\kappa = 1$ of the family (21). However, the history preservation of Proposition 5.1 fails (in the variant $\mathcal{H}_n$ grows even at $c_n = 0$), and the bound R_q of Theorem 5.4 as well as the numerical critical values and trajectories of Section 6 change.

For example, the hierarchy

$M^{(0)}$: "I must not let them notice my tension,"

$M^{(1)}$: "I want to punish the self that tries to hide,"

$M^{(2)}$: "the punishing self is violent, so it too must be punished,"

$M^{(3)}$: "the self that cannot punish correctly must be monitored further"

may arise. Here $\mathcal{O}_0, \mathcal{O}_1, \ldots$ differ in general, because the results of each stage accumulate into $\mathcal{H}_n$ and thereby update the next punishment, justification, and monitoring rules themselves.

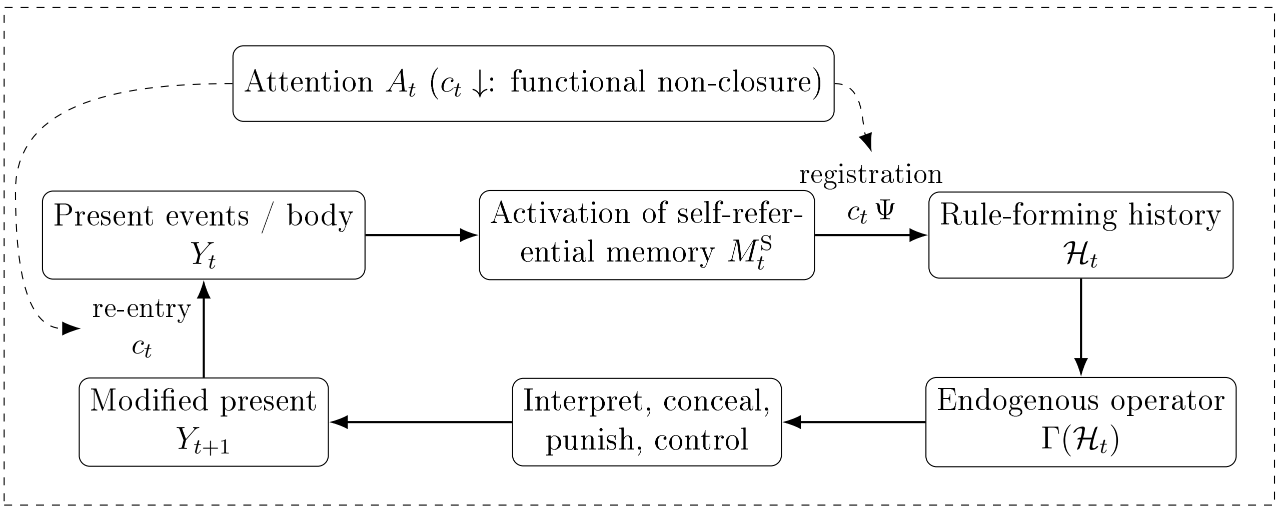
Writing the one-step transition operator as $\Xi_j := D + c_j\,\mathcal{O}_j$, a fixed operator Ξ would give $M^{(n)} = \Xi^n[M^{(0)}]$. In the present model, however,

$$M^{(n)} = (\Xi_{n-1} \circ \cdots \circ \Xi_0)[M^{(0)}], \tag{9}$$

and in general

$$\mathcal{O}_i \circ \mathcal{O}_j \neq \mathcal{O}_j \circ \mathcal{O}_i, \tag{10}$$

so the system is noncommutative and history dependent: outcomes depend on the order of the history. What this noncommutativity actually changes is exhibited by a minimal numerical example in Section 6.4.



Self-referential closed loop (two closure channels)

Figure 1: The self-referential loop: activation is registered into the history (rule-forming channel), and the history generates operators that re-intervene in the present (action channel). Attention (non-closure) weakens both closure channels. Registration into the non-recursive record $\mathcal{K}_t$ is not subject to closure and is omitted from the figure.

4.2 Generation of complexity

The complexity of the model does not arise merely from the number of layers. At each stage, punishment, concealment, justification, and control branch, so the actual structure is closer to a generative tree than to a single series. For example,

$$\mathsf{H}(\mathsf{P}(M)), \quad \mathsf{J}(\mathsf{P}(M)), \quad \mathsf{P}(\mathsf{J}(\mathsf{P}(M))) \tag{11}$$

can be generated simultaneously. Memory endogenously produces not only the defendant, the prosecutor, the defense counsel, and the judge, but also the internal auditor who audits the judge. If one then tries to "resolve" the terms one at a time, the resolution effort creates new terms, so that in principle an infinite regress arises.

4.3 A minimal continuous-time model

Let $x(t)$ be the intensity of recursive memory activity and $p(t)$ the punishment/control gain generated by memory, and consider the minimal system

$$\dot{x}(t) = \{-d + c(t)p(t)\}x(t), \tag{12}$$

$$\dot{p}(t) = -\beta p(t) + \alpha\, c(t)\, x(t), \tag{13}$$

where $d > 0$ is natural decay, $\alpha > 0$ is the sensitivity with which distress activates control operators, and $\beta > 0$ is the decay rate of operator gain. That the closure coefficient also multiplies the rule-forming channel corresponds to (7) in the discrete model (the variant $\dot{p} = -\beta p + \alpha x$, in which only the action channel is gated, does not essentially change the conclusions below). Under complete non-closure $c(t) \equiv 0$, $x(t) = x(0)e^{-dt}$ and $p(t) = p(0)e^{-\beta t}$: distress decays naturally, and the control gain already generated loses its coupling to the present and fades. Here again what non-closure guarantees is the severing of the coupling; the subsequent decay is due to the natural-dynamics assumptions $d > 0$ and $\beta > 0$.

In this system the positive feedback

$$x \uparrow \Rightarrow p \uparrow \Rightarrow c\,p\,x \uparrow \tag{14}$$

is possible. A closed loop, however, is necessary but not sufficient for positive feedback: self-amplification arises only when the effective gain $c\,p$ exceeds the natural decay d.

5 Analytical properties

Proposition 5.1 (The non-closure limit)**.** *In* (6)–(8)*, if $c_n = 0$ for all n, then*

$$M^{(n)} = D^n[M^{(0)}], \qquad \mathcal{H}_n = \mathcal{H}_0, \qquad \mathcal{K}_{n+1} = \mathcal{K}_n + \Theta\big(D^{n+1}[M^{(0)}]\big) \quad (\forall n).$$

That is, rule formation stops completely (the history $\mathcal{H}$ is preserved), registration into the non-recursive record $\mathcal{K}$ continues, and the activation follows the natural dynamics D alone, without operator intervention.

Proof. Substitute $c_n = 0$ and use induction. □

The statement of Section 3—attention does not erase memory but only suspends its sovereignty—is formalized by this proposition. Note that the universal contribution of attention extends only to *preventing self-referential re-amplification and rule formation from being established.* What happens to the active content thereafter depends on the natural dynamics: if the spectral radius $\rho(D) < 1$, the content fades. If $\rho(D) = 1$, then depending on the spectral structure of D (power boundedness), decay, bounded persistence, polynomial growth, and other behaviors can occur. If $\rho(D) > 1$, there exist initial states that grow even when the closed loop is severed (the base system has unstable directions). The phenomenological description "memory appears, yet is not chosen" corresponds to this general case, in which content need not vanish immediately. In an older formulation, where c_n multiplied the entire generated term, lowering c acted as a uniform shrinkage of content and was formally indistinguishable from partial suppression; in the present formulation, lowering c is not a reduction of content but the severing of the closure channels. In the convergence estimates of this section we assume $\delta = \|D\| < 1$ (natural decay).

Proposition 5.2 (Upper bound for the recursive series)**.** *In* (6)*, if the* $\mathcal{O}_n$ *are bounded linear operators, then*

$$\left\|M^{(n)}\right\| \leq \left(\prod_{j=0}^{n-1} \left(\delta + c_j \left\|\mathcal{O}_j\right\|\right)\right) \left\|M^{(0)}\right\|. \tag{15}$$

Proof. Apply the triangle inequality and the submultiplicativity of the operator norm successively to (6) (see Appendix A). □

Corollary 5.3 (Convergence under sufficient non-closure)**.** *If for some* $q < 1$ *and all* n

$$\delta + c_n \left\|\mathcal{O}_n\right\| \leq q, \tag{16}$$

then $M^{(n)} \to 0$ *and* $\sum_{n=0}^{\infty} \left\|M^{(n)}\right\|$ *converges.*

This result shows that attention need not erase each memory content one by one: if the effective gain of every generated operator stays uniformly below 1, the recursive hierarchy decays. The condition of Corollary 5.3, however, is imposed along the trajectory, since $\|\mathcal{O}_n\| = \|\Gamma(\mathcal{H}_n)\|$ itself depends on the solution. The next theorem replaces this assumption by a primitive condition involving only the initial data and the generating map Γ.

Theorem 5.4 (Convergence from a primitive condition)**.** *Suppose* Γ *maps bounded sets to bounded sets. Set* $\Phi(R) := \sup\{\|\Gamma(Y)\| : \|Y\| \leq R\}$ *and* $\bar{c} := \sup_n c_n$*. If some* $q \in (\delta, 1)$ *satisfies*

$$\delta + \bar{c}\,\Phi(R_q) \;\leq\; q, \qquad R_q := \|\mathcal{H}_0\| + \frac{\bar{c}\,q}{1-q} \left\|M^{(0)}\right\|, \tag{17}$$

then $\left\|M^{(n)}\right\| \leq q^n \left\|M^{(0)}\right\|$ *and* $\|\mathcal{H}_n\| \leq R_q$ *for all* n*, and* $\mathcal{H}_n$ *converges in norm.*

Proof. Both estimates are proved simultaneously by induction. Assuming them for $j \leq n$, we have $\|\mathcal{H}_n\| \leq R_q$, hence $\|\Gamma(\mathcal{H}_n)\| \leq \Phi(R_q)$, and (6) gives

$$\left\|M^{(n+1)}\right\| \leq \left(\delta + \bar{c}\,\Phi(R_q)\right) \left\|M^{(n)}\right\| \leq q^{n+1} \left\|M^{(0)}\right\|.$$

Also, by (7) and $\|\Psi(x)\| \leq \|x\|$,

$$\|\mathcal{H}_{n+1}\| \leq \|\mathcal{H}_0\| + \bar{c} \sum_{j=1}^{\infty} \left\|\Psi(M^{(j)})\right\| \leq \|\mathcal{H}_0\| + \bar{c} \sum_{j=1}^{\infty} q^j \left\|M^{(0)}\right\| = R_q.$$

The induction closes. Finally, $\|\mathcal{H}_{n+p} - \mathcal{H}_n\| \leq \bar{c} \left\|M^{(0)}\right\| \sum_{j>n} q^j \to 0$, so $\{\mathcal{H}_n\}$ is Cauchy and converges by completeness. □

The hypothesis of Theorem 5.4 is a checkable condition on the initial data $(M^{(0)}, \mathcal{H}_0)$ and the properties of Γ alone, not on the behavior of the solution. If Γ is given, as in (3), by a linear combination of finitely many basis operators with bounded weight functions, then Φ is bounded by a constant W determined by the operator norms and the weight bounds, and the condition reduces to the inequality $\delta + \bar{c}W \leq q$.

Remark 5.5 (The problem with positing attention as a superior operator). One could instead posit attention as a superior operator generated from memory states, $\mathcal{A}_n = \Gamma_A(\mathcal{H}_n)$, suppressing the control operator through $\mathcal{A}_n \circ \mathcal{O}_n$; hierarchical precision-control architectures in active inference are of exactly this kind, with attentional states modulating the level below and meta-awareness states modulating the level above (Sandved-Smith et al., 2021). But since $\mathcal{A}_n = \Gamma_A(\mathcal{H}_n)$, the attention operator is a function of $\mathcal{H}_n$ and its effects enter the next cumulative history. $\mathcal{A}_n$ is therefore not a privileged controller outside the system but an endogenous operator generated from the in-system state $\mathcal{H}_n$, and the possibility of generating yet another operator that monitors and corrects $\mathcal{A}_n$ remains. Defining attention as merely higher-order control does not dissolve the regress. This remark does not require positing attention as a mysterious entity; what the model requires is only that attention be distinguished not as "a further processing" but as "the condition that processing outputs are not re-entered as the present."

6 Numerical illustrations

6.1 Operator gain depending on cumulative history

Let m_n be the intensity of the recursive layer and h_n its accumulation, and consider the one-dimensional system

$$m_{n+1} = \left\{\delta + c\left(g_0 + g_1 \frac{h_n}{1+h_n}\right)\right\} m_n, \tag{18}$$

$$h_{n+1} = h_n + c\, m_{n+1}. \tag{19}$$

Here $c \in [0, 1]$ is the degree of closure held constant ($c_n \equiv c$), $\delta \in (0, 1)$ is the natural persistence rate, $g_0 > 0$ is the base gain of the rules, $g_1 > 0$ is the rule reinforcement by cumulative history, and the saturation factor $h/(1 + h)$ expresses that reinforcement has a ceiling. The larger h_n, the stronger the rules of punishment, monitoring, and control. Under complete closure $c = 1$, the gain in (18) coincides with the combined gain $\rho_0 + \eta\, h/(1 + h)$ with $\rho_0 := \delta + g_0$ and $\eta := g_1$.

This system is the **one-dimensional specialization** of the general model of Section 4 obtained by taking $\mathscr{M} = \mathbb{R}$, $\mathscr{M}_+ = \mathbb{R}_+$, $D = \delta$, $\Gamma(h) = g_0 + g_1\, h/(1+h)$, and $\Psi =$ the identity (the record $\mathcal{K}$ is omitted in this section); in this case the reduction is exact. Below we assume $\delta + g_0 < 1$: without reinforcement by history, the effective gain stays below 1 (the base gain of the rules alone does not tip the system into amplification). The assumption is used in the case analysis of Section 6.3.

Proposition 6.1 (Closure threshold: a sufficient condition for convergence)**.** *In* (18)–(19)*, if*

$$c \;<\; c^* := \frac{1-\delta}{g_0 + g_1}, \tag{20}$$

then $m_n \leq \{\delta + c(g_0 + g_1)\}^n m_0 \to 0$ and h_n converges to a finite value.

Proof. Since $h/(1+h) < 1$ for all $h \geq 0$, $m_{n+1} < \{\delta + c(g_0 + g_1)\}m_n$. By (20), m_n is dominated by a geometric series, and the sum of the increments of h_n converges. $\square$

6.2 Illustrative results

Take $m_0 = h_0 = 1$, $\delta = 0.30$, $g_0 = 0.35$, $g_1 = 0.55$ (so that $c^* = 0.7/0.9 \approx 0.778$). Under complete functional closure $c = 1$, the initial gain is 0.925 and the response first decays; but once the operator gain, reinforced by the cumulative history, exceeds 1, the system turns to self-amplification. If $c = 0.5 < c^*$, the effective gain is at most 0.75, the recursive layer decays, and the accumulation stays finite. At $c = 0$ (complete non-closure) rule formation stops, and $m_n = \delta^n$ fades with $h_n \equiv h_0$. The last column of Table 1 is the numerical expression of "attention does not erase memory" (Section 3): what is preserved is the history, and what is lost is only the sovereignty of the rules.

Table 1: Evolution of the recursive layer for different degrees of closure ($\delta = 0.30$, $g_0 = 0.35$, $g_1 = 0.55$).

	$c = 1.0$		$c = 0.5$		$c = 0.0$	
n	m_n	h_n	m_n	h_n	m_n	h_n
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	0.9250	1.9250	0.6125	1.3062	0.3000	1.0000
2	0.9361	2.8611	0.3863	1.4994	0.0900	1.0000
4	1.0757	4.9267	0.1595	1.7028	0.0081	1.0000
6	1.3372	7.4549	0.0672	1.7881	0.0007	1.0000
8	1.7375	10.7100	0.0286	1.8243	0.0001	1.0000
10	2.3237	15.0369	0.0122	1.8397	0.0000	1.0000

The numerical example is not intended to reproduce empirical facts; it exhibits a structure. Control effort that looks effective in the short run can turn, with delay, into positive feedback when its results reinforce the next control rules. And attention can move the entire series into the convergent regime, not by inverting the content of the operators, but by weakening the re-entry of their outputs (Figure 2).

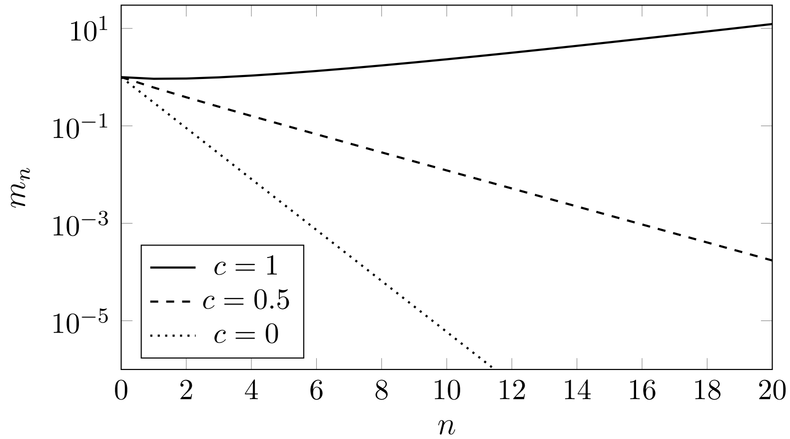


Figure 2: Activation trajectories m_n under different degrees of closure (logarithmic scale; $\delta = 0.30,\ g_0 = 0.35,\ g_1 = 0.55,\ m_0 = h_0 = 1$). Under complete closure $c = 1$ the trajectory turns from initial decay to self-amplification; at $c = 0.5 < c^*$ it decays geometrically; at $c = 0$ it decays purely at the natural persistence rate δ.

6.3 The dichotomy between quieting and self-amplification

Proposition 6.1 is only a sufficient condition. To determine what happens for $c \geq c^*$, consider the family containing (18)–(19),

$$m_{n+1} = G(h_n)\, m_n, \qquad h_{n+1} = h_n + \kappa\, m_{n+1}, \qquad G(h) = a + b\,\frac{h}{1+h} \tag{21}$$

($a > 0,\ b > 0,\ \kappa > 0,\ m_0 > 0,\ h_0 \geq 0$). System (18)–(19) is the case $a = \delta + c\,g_0$, $b = c\,g_1$, $\kappa = c$; the combined-gain form is the case $a = c\rho_0$, $b = c\eta$, $\kappa = 1$. The sequence h_n is strictly increasing, and G is strictly increasing in h.

Theorem 6.2 (Trichotomy and dichotomy)**.** *For the family* (21)*:*

(i) If $a + b \leq 1$, then $m_n \to 0$. If moreover $a + b < 1$, then $m_n \leq (a+b)^n m_0$ and h_n converges to $h_\infty \leq h_0 + \kappa m_0 (a+b)/(1-(a+b))$ (uniform quieting).

(ii) If $a \geq 1$, then $m_n \to \infty$ (uniform amplification).

(iii) If $a < 1 < a + b$ (the critical regime), set

$$h^* := \frac{1-a}{a+b-1} > 0. \tag{22}$$

Then exactly one of the following holds.

(Q) $h_n < h^*$ for all n. Then m_n decreases monotonically to 0, and h_n converges to $h_\infty \leq h^*$ (quieting).

(R) $h_N \geq h^*$ for some N. Then $m_n \to \infty$, with growth rate approaching $a + b$ (persistent self-amplification).

Proof. (i) If $a + b < 1$, then $G(h) < a + b$ gives geometric decay, and the sum of increments of h_n is finite. On the boundary $a+b = 1$, $G(h) < 1$ makes m_n monotonically decreasing with $m_k \le m_0$, so $h_j \le h_0 + \kappa m_0 j$. Then $1 - G(h_j) = b/(1 + h_j) \ge b/(1 + h_0 + \kappa m_0 j)$ sums like a harmonic series, so $\log \prod_j G(h_j) \le -\sum_j (1 - G(h_j)) \to -\infty$, i.e., $m_n \to 0$. (ii) $m_1 \ge a\, m_0 > 0$ gives $h_1 > 0$, and for $n \ge 1$, $G(h_n) \ge a + b\, h_1/(1 + h_1) > 1$, so growth is geometric. (iii) Since h_n is strictly increasing, (Q) and (R) are mutually exclusive and exhaustive. Under (Q), $G(h_n) < G(h^*) = 1$ makes m_n monotonically decreasing; if the limit were positive, the increments κm_{n+1} would be bounded below and $h_n \to \infty$, contradicting (Q); hence $m_n \to 0$ and h_n, bounded and monotone, converges. Under (R), if $h_N > h^*$ then $G(h_n) \ge G(h_N) > 1$ for $n \ge N$ and growth is geometric (if exactly $h_N = h^*$, then $m_{N+1} = m_N > 0$ gives $h_{N+1} > h^*$ and the previous case applies). Then $h_n \to \infty$, so $G(h_n) \to a + b$. □

Proposition 6.1 is a special case of (i). Under the assumption $\delta + g_0 < 1$ we have $a = \delta + c\, g_0 < 1$ for $c \le 1$, so (ii) does not occur, and the substantive division is between uniform quieting and the critical regime, in which **even under identical parameters and an identical degree of closure, the qualitative outcome divides according to whether the cumulative history crosses the critical value h^***. Which side is reached depends on the initial condition and the history. The divergence $m_n \to \infty$ in (R) does not mean that a psychological quantity literally reaches infinity. It is a divergence of the minimal model with saturation omitted; "persistent self-amplification" is used in this paper as a technical shorthand for "entering a regime of sustained amplification beyond the model's range of validity."

Proposition 6.3 (Critical initial value)**.** *In the critical regime with $h_0 < h^*$ fixed, the trajectory is componentwise monotone nondecreasing in m_0, and there exists a critical value $m_0^* \in (0, \infty)$ such that $m_0 \le m_0^*$ implies (Q) and $m_0 > m_0^*$ implies (R). On the boundary trajectory $m_0 = m_0^*$, $h_n \uparrow h^*$. Moreover $m_0^* \le (h^* - h_0)/(\kappa a)$ (proof in Appendix A).*

Corollary 6.4 (Explicit estimate of the quieting basin)**.** *In the critical regime with $h_0 < h^*$, if some $h^\dagger \in (h_0, h^*)$ satisfies*

$$h_0 + \kappa\, m_0 \frac{G(h^\dagger)}{1 - G(h^\dagger)} \;\le\; h^\dagger, \tag{23}$$

then (Q) holds and $m_n \le G(h^\dagger)^n m_0$ (proof in Appendix A).

The case of complete closure $c = 1$ ($a = 0.65,\ b = 0.55,\ \kappa = 1$) lies in the critical regime, with $a + b = 1.2 > 1 > a$ and, for $h_0 = 1$, $h^* = 0.35/0.20 = 1.75$. As Table 2 shows, the value $m_0 = 1$ used in Table 1 is an amplification-side initial condition, crossing at the first step ($h_1 = 1.925 > h^*$), whereas $m_0 = 0.001$ lies on the quieting side. Bisection gives the critical initial value $m_0^* \approx 0.0268$. Lowering the closure to $c = 0.9$, slightly above c^* ($a = 0.615,\ b = 0.495,\ \kappa = 0.9,\ h^* = 3.5$), yields Table 3, with critical initial value $m_0^* \approx 0.1580$.

Table 2: Initial-value dependence under complete closure $c = 1$ ($a = 0.65,\ b = 0.55,\ \kappa = 1,\ h_0 = 1$; $h^* = 1.75$, $m_0^* \approx 0.0268$).

m_0	Outcome	Notes
0.001	Quieting (Q)	$h_\infty \approx 1.0125$
0.02	Quieting (Q)	$h_\infty \approx 1.3495$
0.03	Divergence (R)	h_n crosses h^* at $n = 110$
0.04	Divergence (R)	h_n crosses h^* at $n = 41$
0.05	Divergence (R)	h_n crosses h^* at $n = 27$
0.1	Divergence (R)	h_n crosses h^* at $n = 10$
1	Divergence (R)	h_n crosses h^* at $n = 1$

Table 3: Initial-value dependence at $c = 0.9 > c^*$ ($a = 0.615,\ b = 0.495,\ \kappa = 0.9,\ h_0 = 1$; $h^* = 3.5$, $m_0^* \approx 0.1580$).

m_0	Outcome	Notes
0.05	Quieting (Q)	$h_\infty \approx 1.3288$
0.1	Quieting (Q)	$h_\infty \approx 1.8078$
0.3	Divergence (R)	h_n crosses h^* at $n = 17$
1	Divergence (R)	h_n crosses h^* at $n = 4$

Three observations are in order. First, under identical parameters, a slight difference in initial intensity ($m_0 = 0.02$ versus 0.03) separates quieting from persistent self-amplification. Second, the trajectory with $m_0 = 0.03$ stays below h^* for 110 steps before crossing: amplification after a long incubation—relapse after apparent quiet—can occur without any change in external conditions. Third, lowering the closure from $c = 1$ to $c = 0.9$ does not remove the critical regime, but it greatly enlarges the basin of quieting (m_0^* grows about sixfold). The effect of attention appears not in making divergence impossible but in raising the historical threshold at which self-amplification begins. Figure 3 shows the trajectories as a whole.

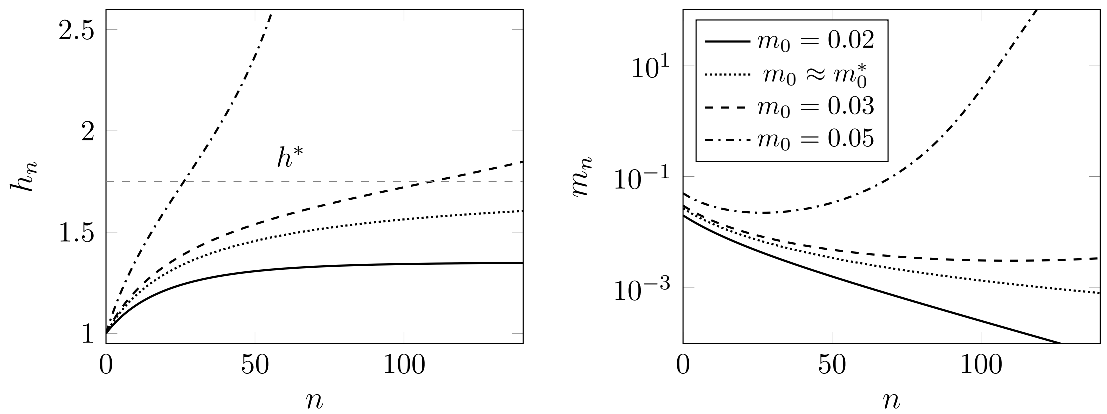


Figure 3: Initial-value dependence under complete closure $c = 1$ ($a = 0.65$, $b = 0.55$, $\kappa = 1$, $h_0 = 1$). Left: cumulative history h_n. Trajectories staying below the critical value $h^* = 1.75$ (gray dashed line) quiet down; trajectories crossing it self-amplify. The trajectory with $m_0 = 0.03$ crosses after an incubation of 110 steps. Right: activation m_n (logarithmic scale); even the amplifying trajectories pass through a long decay phase before turning upward. The dotted curve is a numerical approximation of the separatrix ($m_0 \approx 0.026807$); within the displayed range it is visually indistinguishable from the true boundary trajectory (for which $h_n \uparrow h^*$ by Proposition 6.3).

This history dependence connects with the kindling/stress-sensitization hypotheses in mood-disorder research (Post, 1992; Monroe and Harkness, 2005). The sensitization dynamic—accumulated early episodes lower the threshold for later episodes—is expressed in the present model as the approach of the accumulation h to its critical value h^* (testable forms of the prediction are given in Section 8).

Remark 6.5 (Correspondence with the continuous-time model). In the continuous model (12)–(13) with constant $c > 0$, the origin is a stable node with eigenvalues $-d, -\beta$, and at the interior equilibrium $(x^*, p^*) = (\beta d/(\alpha c^2),\ d/c)$ the Jacobian

$$J^* = \begin{pmatrix} 0 & c\,x^* \\ \alpha c & -\beta \end{pmatrix}, \qquad \det J^* = -\alpha c^2 x^* = -\beta d < 0$$

shows that the interior equilibrium is a saddle (the variant gating only the action channel leaves $\det J^* = -\beta d < 0$ unchanged). The stable manifold of the saddle is the separatrix dividing the state space into quieting and amplification regions, the continuous-time counterpart of the dichotomy in Theorem 6.2(iii).

6.4 Noncommutativity and history dependence of operators: a minimal example

What the noncommutativity of operator composition, noted in Section 4, actually changes is fixed here by a minimal numerical example. Let $\mathscr{M} = \mathbb{R}^2$ and $M = (m_{\mathrm{o}}, m_\ell)^\top$, with m_{o} the overt component (tension and distress present in awareness) and m_ℓ the latent component (load concealed and held outside awareness). Take the basis operators

$$\mathsf{H} = \begin{pmatrix} 0.4 & 0 \\ 0.6 & 1 \end{pmatrix}, \qquad \mathsf{P} = \begin{pmatrix} 1.3 & 0.5 \\ 0 & 1 \end{pmatrix}. \tag{24}$$

H (concealment) transfers 60% of the overt component to the latent one; P (punishment) amplifies overt distress by 30% and recovers latent load into overt distress at rate 0.5—an expression of "even what was hidden becomes material for self-reproach." Then

$$\mathsf{PH} = \begin{pmatrix} 0.82 & 0.5 \\ 0.6 & 1 \end{pmatrix}, \qquad \mathsf{HP} = \begin{pmatrix} 0.52 & 0.2 \\ 0.78 & 1.3 \end{pmatrix}, \qquad [\mathsf{P}, \mathsf{H}] = \begin{pmatrix} 0.3 & 0.3 \\ -0.18 & -0.3 \end{pmatrix} \neq O, \quad (25)$$

and the action on the purely overt state $M^{(0)} = (1, 0)^\top$ divides: $\mathsf{PH}\, M^{(0)} = (0.82,\, 0.6)^\top$ versus $\mathsf{HP}\, M^{(0)} = (0.52,\, 0.78)^\top$. Under "conceal, then punish," the load that was hidden is recovered by punishment into overt distress, which stays high. Under "punish, then conceal," the amplified distress is sent into latency: the surface is calm but the latent load is maximal—a minimal model of the apparently composed but internally pressurized state.

To be precise, since H is invertible in this example, PH and HP are similar, with common eigenvalues $\lambda_1 \approx 1.465$ and $\lambda_2 \approx 0.355$; if one merely iterates the same composite, the asymptotic growth rate does not depend on the order. Noncommutativity produces substantive path dependence because differences in the state profile (the overt/latent allocation) change, through the state-dependent weights $w_k(M)$ of (3), the very operator that is generated and selected next. To exhibit this in minimal form, take as a proxy for Γ the selection rule "if $m_o \geq m_\ell$ apply H (overt dominance triggers the concealment urge), otherwise apply P (latent dominance triggers self-punishment)," and iterate.

Table 4: Six steps under state-dependent selection. Different initial states cause the very sequence of applied operators to diverge.

	$M^{(0)} = (1, 0)^\top$		$M^{(0)} = (0, 1)^\top$	
n	Applied	(m_o, m_ℓ)	Applied	(m_o, m_ℓ)
1	H	(0.4000, 0.6000)	P	(0.5000, 1.0000)
2	P	(0.8200, 0.6000)	P	(1.1500, 1.0000)
3	H	(0.3280, 1.0920)	H	(0.4600, 1.6900)
4	P	(0.9724, 1.0920)	P	(1.4430, 1.6900)
5	P	(1.8101, 1.0920)	P	(2.7209, 1.6900)
6	H	(0.7240, 2.1781)	H	(1.0884, 3.3225)

As Table 4 shows, the system started from pure overtness enters HPHPPH (an alternating conceal–blame cycle), whereas the system started from pure latency enters PPHPPH (sustained self-blame followed by concealment). These strings list the order of application from left to right as a time series; they are not the usual right-to-left notation for operator products. This is a minimal, checkable instance of the paper's claim that the sequence of applied operators is itself a function of the history. Both sequences grow in total magnitude (the spectral radius of the composites exceeds 1); this corresponds to complete closure $c = 1$, and within the framework of (6) the quieting mechanism of the previous subsection operates as c is lowered.

7 Interpretation through everyday examples

7.1 Tension before a class

Suppose bodily tension arises on the way to the classroom. Let the primary state be

$$M^{(0)} = \text{“there is tension in the body.”} \tag{26}$$

At this stage, the perception of a bodily state can be functional. Once past teacher-images and others' evaluations become active, however,

$$M^{(1)} = \text{“the students must not notice”} \tag{27}$$

is generated. The concealment operator H then controls facial expression, voice, and posture, increasing bodily constraint; the increased constraint is re-entered as evidence that "it does look unnatural after all." Next,

$$M^{(2)} = \text{“an I that cannot control even this much tension is no good”} \tag{28}$$

is added, and the punishment operator P engages.

The problem is then no longer only the primary state of tension. The operators generated to process the primary state reinforce, through the results of their operation, both the primary state and themselves. Yet when tension, the urge to hide, bodily control, and self-punishment are seen together as one movement, the operator outputs become less liable to be re-entered as the present itself, even if none of them vanishes at once. The model represents this mode as $c \downarrow$.

7.2 Contrast with mathematical curiosity

When one does mathematics "because it is interesting," memory works massively: definitions, theorems, past failures, and calculational rules are consulted, and hypotheses are generated. But here the present mathematical problem is using memory as a tool; the mathematics is not necessarily being used to maintain "an I that can prove things."

The paper therefore does not classify the mere presence of thought, purpose, or self-mention as psychological memory. The dividing question is whether the task is using memory, or memory is using the task for the maintenance of a self-image. The same mathematical activity can shift from functional to psychological as its center of gravity moves from curiosity about the result to self-evaluation.

8 Implications for psychiatry and clinical psychology

8.1 From personal failure to dynamic formulation

Patients often understand themselves as "unable to control anxiety," "unable to change their thinking," "unable to carry out the treatment properly," and their self-blame deepens. In the present model, one considers the possibility that, in response to primary distress, control operators are generated from the same memory system and their operation forms a positive feedback. The problem is then not only an insufficiency of an independent self's control capacity, but the fact that the controller is built into the pathological dynamics as an endogenous operator generated from the same memory state.

This reformulation is not meant to absolve responsibility. Its purpose is to move the character judgment "I am weak" into an observable, shareable, testable dynamics—an object the patient and the therapist can work on together.

8.2 The possibility that treatment effort creates new operators

Therapeutic instructions can create new self-referential operators of the form

> I must become more attentive. I must eliminate the anxiety. I must observe correctly.

The evaluation of a treatment should therefore examine not only symptom change but whether the technique is generating new monitors and punishers inside the patient.

That said, functional cognitive control, safety behaviors in the proper sense, medication adherence, sleep regulation, and impulse control can be indispensable. The paper does not reject control in general. What is hypothesized to be harmful is self-referential control: the repeated monitoring of the gap between self-image and ideal state, and the forced, ongoing correction of inner experience.

Nor should non-closure be confused with reduced contact with inner experience. Non-closure means none of dissociation or depersonalization, emotional numbing, abandonment of reality testing, or avoidance of necessary action. It refers to retaining contact with the content and reality testing while the content does not govern the present as a self-referential rule. The distinction is of the same type as that drawn in decentering research, which separates decentering from detachment and avoidance and organizes it as multiple processes (Bernstein et al., 2015).

8.3 Testing with within-person time series

The model is testable with experience-sampling methods and smartphone-based within-day records. For instance:

$$x_t \uparrow \Rightarrow p_{t+1} \uparrow, \tag{29}$$

$$p_t \uparrow \Rightarrow x_{t+1} \uparrow \quad \text{(controlling for current } x_t\text{)}, \tag{30}$$

$$A_t \uparrow \Rightarrow \text{the coupling } p_t \to x_{t+1} \text{ weakens}, \tag{31}$$

where x_t is primary distress, p_t is concealment/punishment/correction effort, and A_t is an observable index close to decentering or nonreactive awareness. Dynamic structural equation models, person-specific state-space models, and network time-series analysis can probe the within-person bidirectionality and lags. Beyond coupling strengths, the dichotomy of Section 6.3 predicts a main effect of the cumulative quantity: controlling for the decentering level (the index corresponding to A_t), individuals and periods with a larger accumulation of self-referential episodes should show higher residual rumination and higher persistence probability after provocations of equal intensity. Estimating a proxy of the accumulation together with post-provocation decay rates makes this prediction testable in within-person time series as well. A further, sharper consequence follows from the mechanism itself: as the effective gain $\delta + c\,g(h_n)$ approaches one near the threshold, within-person autocorrelation of the recursive activity should rise—precisely the critical-slowing-down and emotional-inertia signatures reported for mood dynamics (van de Leemput et al., 2014; Kuppens et al., 2010)—so that rising inertia should precede history-driven transitions in this model too.

Hypothesis (31) deserves a caveat. Under the present definition of attention as the first-person experience of low closure, (31) is less a model-specific prediction than a test of construct validity: whether the measure of A_t captures the theoretical closure c_t. The model-specific novel predictions are the bidirectional coupling (29)–(30), the main effect of the cumulative quantity described above, history-dependent divergence of outcomes under identical stimuli, and the raising of the critical initial value m_0^* by lower closure (compare Tables 2 and 3).

8.4 An opening toward positive modes of experience

The core of the paper is the dynamics up to quieting; it does not model what appears after quieting. Whether the phase in which self-referential operators have lost sovereignty connects not merely to symptomlessness or apathy but to positive modes of experience—attention, freedom, stillness, openness of relation—is an important question for future research. If that possibility is imposed as a treatment goal, it generates new operators; it must be treated strictly as "what may appear."

9 A limited correspondence with contemplative descriptions

This paper assumes no religious or metaphysical premises. Still, checking how far an independently constructed model corresponds to older contemplative descriptions has value as triangulation of the phenomenon.

Krishnamurti repeatedly stated that the observer is not separate from the observed: the observer is itself formed from the past, from experience, knowledge, and memory (Krishnamurti, 1973). In the vocabulary of this paper, the controller is not an agent outside the system but an endogenous operator generated by $\Gamma(\mathcal{H}_n)$. And the refusal to posit a further agent controlling the observer from above corresponds to modeling attention not as an operator but as a non-closure condition.

This correspondence does not justify the model by contemplative authority. It merely checks hypotheses obtained from everyday examples and mathematical structure against an independent articulation. Comparisons with Christian, Buddhist, and other religious or philosophical descriptions are left to a separate paper, since they could prematurely activate readers' existing images and obscure the memory mechanism at the core of this paper.

10 Discussion

10.1 Where does the cunning of memory come from?

Memory is not a malicious agent. Yet when the following properties combine, the resulting behavior is remarkably cunning.

(1) It does not display itself as a mapping of the past; it carries the presence of "this is the present."

(2) It generates self-images and goals and detects the gap with the present as error.

(3) It generates operators that correct the error from the same memory.

(4) It re-enters the results of operation as evidence of the operators' legitimacy.

(5) When an operator fails, it generates a superior operator.

(6) It can generate even the self-image of the researcher or winner: "the I that has seen through the whole."

Memory is thus a self-modifying system that simultaneously generates the object, the problem, the method, and the evaluator. Its complexity lies not in the individual thought contents but in the generating rule Γ. The research focus should accordingly be not on erasing each term, but on the conditions under which the effective closure of Γ rises or falls.

10.2 Hyper-complex memory and simple attention

Self-referential memory generates punishers of punishers and proliferates hierarchy and branching. Attention, by contrast, can appear—at least experientially—as a simple event without hierarchy, purpose, comparison, or punisher. "Simple" here does not mean easy; it means without self-referential re-entry.

Importantly, attention need not be the reward for having solved memory's complex computation to the end. When the generating rule as a whole is seen and the loop no longer closes, the need to maintain the complexity is lost. In this sense attention is not the topmost state of a hyper-complex system but the phase in which the jurisdiction of complexification has lapsed.

10.3 Limitations

The paper has at least six limitations. First, memory, attention, and operator are operational theoretical constructs; no one-to-one correspondence with neuroscientific entities is claimed. Second, the boundary between functional and psychological memory is context dependent, and inter-observer reliability needs to be established. Third, representing attention by a closure coefficient drastically compresses the phenomenology of attention. Fourth, since thought suppression and self-control can be adaptive, the sign and time scale of the positive feedback must be identified empirically. Fifth, the numerical examples illustrate theoretical structure and are not fitted to clinical data. Sixth, the pathway by which non-recursive records are later recalled and reused self-referentially, flowing into the history ($\mathcal{K} \to \mathcal{H}$), is omitted; handling the recruitment of episodic memory by rumination requires introducing this coupling.

11 Conclusion

This paper has described the complication of psychological distress not only as growth of memory contents but as a recursive structure in which memory endogenously generates and updates the operators that process memory itself. In functional memory, the present task uses memory as a tool. In self-referential psychological memory, memory evaluates and controls the present through a self-image and re-enters the outcome as evidence of its own correctness.

In this structure the controller is not independent of the controlled. The memory that conceals, the memory that punishes it, and the memory that monitors the punisher are generated from the same system. Consequently, controlling the hierarchy term by term can add a new level and raise the order of the problem. The paper represented this structure by the map Γ from memory states to operators and by history-dependent operator composition.

Attention was formalized not as a superior control operator but as functional non-closure: the condition that outputs of memory operators are not re-entered as present facts. In the non-closure limit, memory is not erased: the rule-forming history is preserved, registration into non-recursive records continues, and what attention guarantees extends only to stopping self-referential re-amplification and rule formation, the fate of the active content being left to the natural dynamics. The analysis and the numerical illustrations showed, further, that sufficient non-closure guarantees convergence of the recursive series under a checkable primitive condition on the initial data and the generating map, whereas in the critical regime quieting and persistent self-amplification are divided by an explicit critical value of the cumulative history, so that identical conditions can lead to different outcomes depending on history. The possibility of self-amplification after a long incubation, and the implication that improved attention raises the historical threshold for amplification rather than making divergence impossible, connect with kindling and stress-sensitization research on mood disorders. As for noncommutativity, a minimal example confirmed that exchanging the order of application changes the state profile and, through the state-dependent generating rule, bifurcates the operator sequence itself.

The model moves self-blame into dynamic case formulation, examines the possibility that treatment effort itself generates new control operators, and yields predictions falsifiable with within-person time series. It also leaves open the possibility that the mechanical movement of memory is not the whole of the human being. The next tasks are the operationalization of endogenous operators and attention, testing with everyday time series, discrimination between functional control and harmful self-referential control, and the study of positive modes of experience after quieting.

A Supplements to the propositions

A.1 Stepwise expansion for the norm bound (Proposition 5.2)

By (6) and the triangle inequality,

$$\begin{aligned}
\|M^{(1)}\| &\leq (\delta + c_0 \|\mathcal{O}_0\|)\, \|M^{(0)}\|\,, \\
\|M^{(2)}\| &\leq (\delta + c_1 \|\mathcal{O}_1\|)\, \|M^{(1)}\| \\
&\leq (\delta + c_1 \|\mathcal{O}_1\|)(\delta + c_0 \|\mathcal{O}_0\|)\, \|M^{(0)}\|\,.
\end{aligned}$$

Repeating the step yields (15).

A.2 Supplementary proofs for the dichotomy

Proof of Proposition 6.3. If $m_0 \leq m_0'$, monotonicity and positivity of G give, by induction, $m_n \leq m_n'$ and $h_n \leq h_n'$. If m_0' is (Q), then $h_n \leq h_n' < h^*$, so m_0 is (Q). Hence the (Q)-set is downward closed and the (R)-set upward closed; setting $m_0^* := \sup\{m_0 > 0 : \text{(Q)}\}$, every $m_0 < m_0^*$ is (Q) and every $m_0 > m_0^*$ is (R). Finiteness follows from $h_1 \geq h_0 + \kappa a\, m_0$: if $m_0 > (h^* - h_0)/(\kappa a)$, then $h_1 > h^*$ and (R) holds. Positivity follows because the condition of Corollary 6.4 is satisfied for sufficiently small m_0.

We now settle the boundary $m_0 = m_0^*$. Suppose first that the boundary trajectory reaches $h_N \geq h^*$ at some finite time N. If exactly $h_N = h^*$, then $m_{N+1} = G(h^*)m_N = m_N > 0$ gives $h_{N+1} > h^*$, so without loss of generality $h_N > h^*$. The finite-time state (m_N, h_N) is a continuous function of m_0, so for sufficiently small $\varepsilon > 0$ the trajectory with $m_0 = m_0^* - \varepsilon$ also has $h_N > h^*$, i.e., is (R). This contradicts the fact that all of $(0, m_0^*)$ is (Q). Hence the boundary trajectory satisfies $h_n < h^*$ for all n and belongs to (Q); in particular $m_n \to 0$ and h_n increases monotonically to some $h_\infty \leq h^*$.

Suppose next that $h_\infty < h^*$, and pick $h^\dagger$ with $h_\infty < h^\dagger < h^*$. Since $m_N \to 0$ and $h_N \to h_\infty$, for all sufficiently large N

$$h_N + \kappa\, m_N \frac{G(h^\dagger)}{1 - G(h^\dagger)} \;<\; h^\dagger$$

holds with strict inequality. The left-hand side and the finitely many values $h_0, \ldots, h_N$ are continuous in m_0, so for all m_0 in some neighborhood of m_0^*—including values above m_0^*—the same strict inequality holds at time N, and $h_j < h^*$ $(j \leq N)$ is preserved. Applying the induction of Corollary 6.4 with time N as initial time, these trajectories remain at $h_n \leq h^\dagger < h^*$ for $n \geq N$ and are (Q). This contradicts the fact that every $m_0 > m_0^*$ is (R). Hence $h_\infty = h^*$: on the boundary trajectory, $h_n \uparrow h^*$. □

Proof of Corollary 6.4. Set $q := G(h^\dagger) < 1$ and prove $h_n \leq h^\dagger$ and $m_n \leq q^n m_0$ simultaneously by induction. Assuming both, $m_{n+1} = G(h_n)m_n \leq q^{n+1}m_0$, and

$$h_{n+1} = h_0 + \kappa \sum_{j=1}^{n+1} m_j \leq h_0 + \kappa m_0 \sum_{j=1}^{\infty} q^j = h_0 + \kappa m_0 \frac{q}{1-q} \leq h^\dagger.$$

The induction closes, and $h_n \leq h^\dagger < h^*$ for all n. □

A.3 A fixed-point representation

Compressing time and supposing that the total memory M is the sum of the initial memory M_0 and the self-operation, one obtains the nonlinear fixed-point problem

$$M = M_0 + \Gamma(M)[M]. \tag{32}$$

In the approximation with $\Gamma(M)$ frozen, if the spectral radius is below one,

$$M = \{I - \Gamma(M)\}^{-1} M_0 \tag{33}$$

can be written formally. But since Γ itself depends on M, a self-consistent solution is required. This expresses in a single line the central structure of the paper: the rule that tries to solve the problem changes with the problem state.

References

Abramowitz, J. S., Tolin, D. F., & Street, G. P. (2001). Paradoxical effects of thought suppression: A meta-analysis of controlled studies. *Clinical Psychology Review*, 21(5), 683–703. https://doi.org/10.1016/S0272-7358(00)00057-X

Bernstein, A., Hadash, Y., Lichtash, Y., Tanay, G., Shepherd, K., & Fresco, D. M. (2015). Decentering and related constructs: A critical review and metacognitive processes model. *Perspectives on Psychological Science*, 10(5), 599–617. https://doi.org/10.1177/1745691615594577

Bernstein, A., Hadash, Y., & Fresco, D. M. (2019). Metacognitive processes model of decentering: Emerging methods and insights. *Current Opinion in Psychology*, 28, 245–251. https://doi.org/10.1016/j.copsyc.2019.01.019

Bolderston, H., Gillanders, D. T., Turner, G., Taylor, H. C., Ní Mhaoileoin, D., & Coleman, A. (2019). The initial validation of a state version of the Cognitive Fusion Questionnaire. *Journal of Contextual Behavioral Science*, 12, 207–215. https://doi.org/10.1016/j.jcbs.2018.04.002

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, 16(1), 5–13. https://doi.org/10.1002/wps.20375

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, 9, 91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608

Brosschot, J. F., Gerin, W., & Thayer, J. F. (2006). The perseverative cognition hypothesis: A review of worry, prolonged stress-related physiological activation, and health. *Journal of Psychosomatic Research*, 60(2), 113–124. https://doi.org/10.1016/j.jpsychores.2005.06.074

Carver, C. S., & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin*, 92(1), 111–135. https://doi.org/10.1037/0033-2909.92.1.111

Gilbert, P., & Procter, S. (2006). Compassionate mind training for people with high shame and self-criticism: Overview and pilot study of a group therapy approach. *Clinical Psychology & Psychotherapy*, 13(6), 353–379. https://doi.org/10.1002/cpp.507

Gillanders, D. T., Bolderston, H., Bond, F. W., Dempster, M., Flaxman, P. E., Campbell, L., Kerr, S., Tansey, L., Noel, P., Ferenbach, C., Masley, S., Roach, L., Lloyd, J., May, L., Clarke, S., & Remington, B. (2014). The development and initial validation of the Cognitive Fusion Questionnaire. *Behavior Therapy*, 45(1), 83–101. https://doi.org/10.1016/j.beth.2013.09.001

Goldstein, A. J., & Chambless, D. L. (1978). A reanalysis of agoraphobia. *Behavior Therapy*, 9(1), 47–59.

Haslbeck, J. M. B., Ryan, O., Robinaugh, D. J., Waldorp, L. J., & Borsboom, D. (2022). Modeling psychopathology: From data models to formal theories. *Psychological Methods*, 27(6), 930–957. https://doi.org/10.1037/met0000303

Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (1999). *Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change*. New York: Guilford Press.

Higginson, S., Mansell, W., & Wood, A. M. (2011). An integrative mechanistic account of psychological distress, therapeutic change and recovery: The Perceptual Control Theory approach. *Clinical Psychology Review*, 31(2), 249–259. https://doi.org/10.1016/j.cpr.2010.01.005

Koster, E. H. W., De Lissnyder, E., Derakshan, N., & De Raedt, R. (2011). Understanding depressive rumination from a cognitive science perspective: The impaired disengagement hypothesis. *Clinical Psychology Review*, 31(1), 138–145. https://doi.org/10.1016/j.cpr.2010.08.005

Krishnamurti, J. (1973). When the mind realises the observer is the observed there is a release of energy. Public talk, Rome, October 28, 1973. Official transcript, Krishnamurti Foundation. https://www.krishnamurti.org/transcript/when-the-mind-realises-the-observer-is-the-observed-there-is-a-release-of-energy/

Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological Science*, 21(7), 984–991. https://doi.org/10.1177/0956797610372634

Laukkonen, R. E., & Slagter, H. A. (2021). From many to (n)one: Meditation and the plasticity of the predictive mind. *Neuroscience & Biobehavioral Reviews*, 128, 199–217. https://doi.org/10.1016/j.neubiorev.2021.06.021

Lutz, A., Jha, A. P., Dunne, J. D., & Saron, C. D. (2015). Investigating the phenomenological matrix of mindfulness-related practices from a neurocognitive perspective. *American Psychologist*, 70(7), 632–658. https://doi.org/10.1037/a0039585

Monroe, S. M., & Harkness, K. L. (2005). Life stress, the “kindling” hypothesis, and the recurrence of depression: Considerations from a life stress perspective. *Psychological Review*, 112(2), 417–445. https://doi.org/10.1037/0033-295X.112.2.417

Morita, S. (1998). *Morita Therapy and the True Nature of Anxiety-Based Disorders (Shinkeishitsu)* (A. Kondo, Trans.; P. LeVine, Ed.). Albany: State University of New York Press. (Original work published 1928)

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science*, 3(5), 400–424. https://doi.org/10.1111/j.1745-6924.2008.00088.x

Post, R. M. (1992). Transduction of psychosocial stress into the neurobiology of recurrent affective disorder. *American Journal of Psychiatry*, 149(8), 999–1010. https://doi.org/10.1176/ajp.149.8.999

Powers, W. T. (1973). *Behavior: The Control of Perception*. Chicago: Aldine.

Roberts, H., Watkins, E. R., & Wills, A. J. (2013). Cueing an unresolved personal goal causes persistent ruminative self-focus: An experimental evaluation of control theories of rumination. *Journal of Behavior Therapy and Experimental Psychiatry*, 44(4), 449–455. https://doi.org/10.1016/j.jbtep.2013.05.004

Robinaugh, D. J., Haslbeck, J. M. B., Waldorp, L. J., Kossakowski, J. J., Fried, E. I., Millner, A. J., McNally, R. J., Ryan, O., de Ron, J., van der Maas, H. L. J., van Nes, E. H., Scheffer, M., Kendler, K. S., & Borsboom, D. (2024). Advancing the network theory of mental disorders: A computational model of panic disorder. *Psychological Review*, 131(6), 1482–1508. https://doi.org/10.1037/rev0000515

Sandved-Smith, L., Hesp, C., Mattout, J., Friston, K., Lutz, A., & Ramstead, M. J. D. (2021). Towards a computational phenomenology of mental action: Modelling meta-awareness and attentional control with deep parametric active inference. *Neuroscience of Consciousness*, 2021(1), niab018. https://doi.org/10.1093/nc/niab018

Smallwood, J., & Schooler, J. W. (2015). The science of mind wandering: Empirically navigating the stream of consciousness. *Annual Review of Psychology*, 66, 487–518. https://doi.org/10.1146/annurev-psych-010814-015331

Smith, J. M., & Alloy, L. B. (2009). A roadmap to rumination: A review of the definition, assessment, and conceptualization of this multifaceted construct. *Clinical Psychology Review*, 29(2), 116–128. https://doi.org/10.1016/j.cpr.2008.10.003

Suzuki, T. (2026). Two dynamics toward mental quieting: A comparative model of faith-mediated and self-observation pathways [In Japanese]. *The Waseda Commercial Review*, 472, forthcoming. (English version: Working Paper No. 2026-006, Research Institute of Business Administration, Waseda University.)

Teasdale, J. D. (1983). Negative thinking in depression: Cause, effect, or reciprocal relationship? *Advances in Behaviour Research and Therapy*, 5(1), 3–25. https://doi.org/10.1016/0146-6402(83)90013-9

Teasdale, J. D., Moore, R. G., Hayhurst, H., Pope, M., Williams, S., & Segal, Z. V. (2002). Metacognitive awareness and prevention of relapse in depression: Empirical evidence. *Journal of Consulting and Clinical Psychology*, 70(2), 275–287. https://doi.org/10.1037/0022-006X.70.2.275

van de Leemput, I. A., Wichers, M., Cramer, A. O. J., Borsboom, D., Tuerlinckx, F., Kuppens, P., van Nes, E. H., Viechtbauer, W., Giltay, E. J., Aggen, S. H., Derom, C., Jacobs, N., Kendler, K. S., van der Maas, H. L. J., Neale, M. C., Peeters, F., Thiery, E., Zachar, P., & Scheffer, M. (2014). Critical slowing down as early warning for the onset and termination of depression. *Proceedings of the National Academy of Sciences*, 111(1), 87–92. https://doi.org/10.1073/pnas.1312114110

Watkins, E. R. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin*, 134(2), 163–206. https://doi.org/10.1037/0033-2909.134.2.163

Watkins, E. R., & Nolen-Hoeksema, S. (2014). A habit-goal framework of depressive rumination. *Journal of Abnormal Psychology*, 123(1), 24–34. https://doi.org/10.1037/a0035540

Wegner, D. M., Schneider, D. J., Carter, S. R., III, & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology*, 53(1), 5–13. https://doi.org/10.1037/0022-3514.53.1.5

Wells, A. (2009). *Metacognitive Therapy for Anxiety and Depression.* New York: Guilford Press.

Wells, A., & Matthews, G. (1994). *Attention and Emotion: A Clinical Perspective.* Hove, UK: Lawrence Erlbaum Associates.