Attention as Distinction Admission in Finite Systems

Capacity-Limited Gating, Boundary Relevance, and Tunnel Vision

Yining Wu
Independent Researcher
yining.wu@alumni.upenn.edu

Version v1.0 – May 2026

Abstract

This paper develops the attention paper in the agency-semantics spine of Distinction Theory. Building on the formal core of Active Finite Distinction Systems and the agency-semantics spine, it treats attention as capacity-limited admission of candidate distinctions into an update channel. A distinction becomes attended when it is admitted, maintained or made recoverable, assigned a verification path or status, and made available for downstream update, action, prediction, coordination, or boundary maintenance. The paper separates attention from salience: salient distinctions may be rejected when their encoding, verification, or maintenance cost exceeds available capacity, while low-salience distinctions may be admitted when they have high causal boundary relevance or high exploratory information gain. The paper introduces candidate distinction streams, admission gates, scalar and vector attention budgets, boundary-gradient attention value, verification load, background scanning, semantic deficit, nonlinear gate steepening, tunnel vision, attention hysteresis, attention failure modes, collective attention, and audit protocols for biological, cognitive, artificial, organizational, and civilization-scale systems. A deterministic synthetic normal-form model illustrates salience-value dissociation, capacity-limited admission, verification-limited rejection, deficit-induced tunnel vision, hysteretic recovery, and collective attention under epistemic pollution. The paper does not replace cognitive neuroscience, rational inattention, predictive processing, or transformer attention. It supplies a conservative finite-system bridge for analyzing attention as distinction admission under capacity, verification, and resource constraints.

Keywords: attention; distinction admission; finite capacity; boundary maintenance; salience; verification load; tunnel vision; active finite distinction systems; rational inattention; information bottleneck; transformer attention; organizational attention; epistemic pollution.

Epistemic Notice and Scope

This manuscript is an agency-semantics spine paper, not a completed theory of cognitive attention, consciousness, neural competition, or machine attention. It does not claim that all attention is conscious attention. It does not claim that salience equals value, that current transformer attention layers are sufficient for FDS attention, or that attention can be reduced to a single scalar. Its claim is narrower: once a system is modeled as an active finite distinction system with a boundary, memory, update rule, action space, finite capacity, and boundary-maintenance loss, attention can be treated as the capacity-limited admission of candidate distinctions into an update-relevant channel.

The paper follows the layered discipline of the FDS formal core and the agency-semantics spine: formal definitions, normal-form models, bridge assumptions, and domain applications must be separated. A failure of a biological, cognitive, artificial, organizational, or social mapping may demote that mapping without refuting the formal FDS core [1, 2]. The deterministic numerical model is illustrative only. It visualizes the definitions and failure modes; it is not empirical evidence.

1 Introduction

A finite system cannot attend to everything it can detect. Detection produces candidate distinctions; attention admits a subset of them into update. The difference matters. A flash, a word, a warning, a memory, a social signal, a token in a model context, or a weak anomaly in an instrument can be present without becoming operationally available. M1 begins from this gap between availability and admission.

The FDS core defines an active finite distinction system as a tuple

$$S = (X, E, B, M, Y, A, U, \pi, \ell, \Phi, \mathcal{P}, \tau), \tag{1}$$

where X is internal state, E environment, B boundary, M memory/model space, Y observation channel, A action space, U update map, π finite projection, ℓ boundary-maintenance loss, Φ resource budget, $\mathcal{P}$ perturbation/pruning family, and τ update timescale [1]. The M0 agency-semantics paper introduced a dependency skeleton

$$\text{distinction} \to \text{record} \to \text{attention} \to \text{value} \to \text{goal} \to \text{meaning} \to \text{agency} \to \text{culture}, \tag{2}$$

and defined attention as capacity-limited distinction admission into an update channel [2]. M1 expands the first active step in that chain.

1.1 Why attention needs a finite-system formulation

Attention is often described as selection, orienting, priority, competition, or limited-capacity processing. M1 uses a different but compatible abstraction. It asks: when a finite boundary-maintaining system faces more candidate distinctions than it can encode, verify, maintain, and update, which distinctions are admitted to the next state?

This admission view is useful because it applies across scales. A cell admits some chemical gradients into regulatory response. A brain admits some sensory and memory distinctions into task control. A machine system admits some tokens, tool outputs, or observations into its writable state. An organization admits some warnings into agenda and policy. A civilization admits some claims into archives, standards, or science. In each case, the system must choose because its update channel is finite.

1.2 Attention is not salience

The central distinction in this paper is between salience and attention. Salience is a property of a candidate distinction relative to the current perceptual, informational, or social field: it stands out. Attention is admission into an update-relevant channel. A salient distinction can fail attention if it is too costly to verify, too expensive to maintain, irrelevant to the boundary task, or suppressed by a higher-priority gate. Conversely, a low-salience distinction can be attended if it has high causal boundary relevance or high information gain for a model that is becoming brittle.

1.3 What this paper does not claim

This paper does not replace cognitive neuroscience, psychological theories of selective attention, rational inattention, active inference, or transformer attention. It does not say that attention is always optimal. It does not claim that every admission gate is conscious. It does not claim that an attention matrix inside a neural network is attention in the full FDS sense. It supplies a finite-system bridge: attention is constrained admission of distinctions into update.

Table 1: Claim-status summary for M1. The table is an audit device: several entries are formal or operational bridge claims, not established empirical results.

Claim ID	Tier	Claim	Failure or demotion condition
M1-001	Formal bridge	Attention is capacity-limited distinction admission into an update channel.	Attention-like selection occurs without finite capacity, admission, update gating, or priority constraint under the specified mapping.
M1-002	Operational bridge	Salience and attention are separable. Salient distinctions can be rejected if cost or verification burden is too high.	Empirical systems always admit highest-salience items regardless of cost, capacity, task, or verification constraints.
M1-003	Operational bridge	Boundary-efficient or loss-minimizing attention systems preferentially admit high causal boundary-value distinctions under controlled capacity conditions.	Admission patterns are no better predicted by causal boundary value than by raw salience or noise under a valid mapping.
M1-004	Formal / model bridge	Attention allocation can be written as constrained optimization over value, curiosity, cost, and capacity.	No useful mapping exists between admission patterns and constrained allocation variables.
M1-005	Operational bridge	Semantic or attention deficit steepens admission thresholds and can produce tunnel vision.	High load or deficit produces no narrowing, thresholding, or priority collapse in systems claimed to have finite attention.
M1-006	AI / cognition bridge	Artificial attention belongs to a coupled architecture only when routed distinctions affect durable update, action, maintenance, or verification.	Bare attention weights alone satisfy strong FDS attention without durable update or downstream relevance.
M1-007	Social bridge	Collective attention is shared admission under finite communication, verification, and externalized memory capacity.	Group-scale attention shows no relation to verification capacity, agenda-setting, or externalized memory.
M1-008	Failure-mode bridge	Attention failure includes overload, distraction, salience capture, suppression, tunnel vision, false admission, and critical distinction exclusion.	These failure modes cannot be operationalized as admission errors under finite capacity.
M1-009	Operational bridge	Attention recovery after deficit-induced narrowing can lag behind external load reduction because of hysteresis in gate thresholds, verification routines, or maintained threat priors.	Attention gates relax immediately and without lag after load reduction in systems where hysteresis is claimed.

2 FDS and M0 background

2.1 Active boundary relevance

The FDS core applies its maintenance and deficit logic to active-boundary systems rather than to every bounded object. A minimal relevance screen is

$$P\big(U(M_t, Y_t) \neq M_t\big) > 0, \qquad I(M_{t+1}; \ell_{t+k}) > 0 \tag{3}$$

for some $k > 0$. In empirical settings this should be strengthened to an intervention or ablation test:

$$\mathbb{E}[\ell_{t+k} \mid \mathrm{do}(U)] \neq \mathbb{E}[\ell_{t+k} \mid \mathrm{do}(U_\emptyset)]. \tag{4}$$

Attention matters to FDS only when admission affects a future update, record, action, verification, or boundary-maintenance variable.

2.2 Candidate distinctions and update channels

Let $\mathcal{D}_t = \{d_{t,1}, \ldots, d_{t,n}\}$ denote the candidate distinction stream during update window t. Candidate distinctions may arise from sensors, memory, internal simulation, social communication, external archives, tool outputs, or delegated systems. The update channel is the part of the system through which admitted distinctions can affect M_{t+1}, action selection, verification, or future boundary-maintenance loss.

M1 distinguishes three events:

1. *availability*: a distinction exists in the candidate stream;
2. *detection*: the system can register or measure the distinction transiently;
3. *attention*: the distinction is admitted into update-relevant processing.

Detection without admission is possible. Admission without durable maintenance is possible. Reportability is downstream of attention: FDS attention can occur without report, and report can fail after transient admission if maintenance or communication capacity is insufficient.

2.3 Notation alignment

Throughout the paper ℓ_{maint} denotes boundary-maintenance loss. It is not physical entropy production. C_{att} denotes attention admission capacity; C_{verify} denotes verification capacity; C_{sem} denotes maintained semantic capacity. The physical entropy ledger, when relevant, is denoted Σ_{phys} and is not identified with ℓ_{maint}. Admission, verification, and maintenance may have physical costs under bridge assumptions, but M1 does not equate an abstract attention choice with a fixed thermodynamic heat cost.

3 Definition: attention as distinction admission

Definition 1 (Candidate distinction stream)**.** A candidate distinction stream $\mathcal{D}_t$ is the set of possible distinctions available to a finite system during an update window. Candidate distinctions may be external signals, internal memories, simulated alternatives, social claims, environmental affordances, or outputs of delegated systems.

Definition 2 (Attention gate)**.** An attention gate is a map

$$a_t : \mathcal{D}_t \to [0, 1], \tag{5}$$

where $a_t(d)$ is the admission weight assigned to distinction d. Hard attention uses $a_t(d) \in \{0, 1\}$; soft attention uses $a_t(d) \in [0, 1]$.

Definition 3 (FDS attention)**.** A distinction d is attended by system S at time t if it is admitted through an attention gate, encoded into a maintained or update-relevant record, assigned a verification path or verification status, and made available to downstream update, action, prediction, verification, coordination, or boundary maintenance. Attention is capacity-limited admission of candidate distinctions into the update channel of an active finite system.

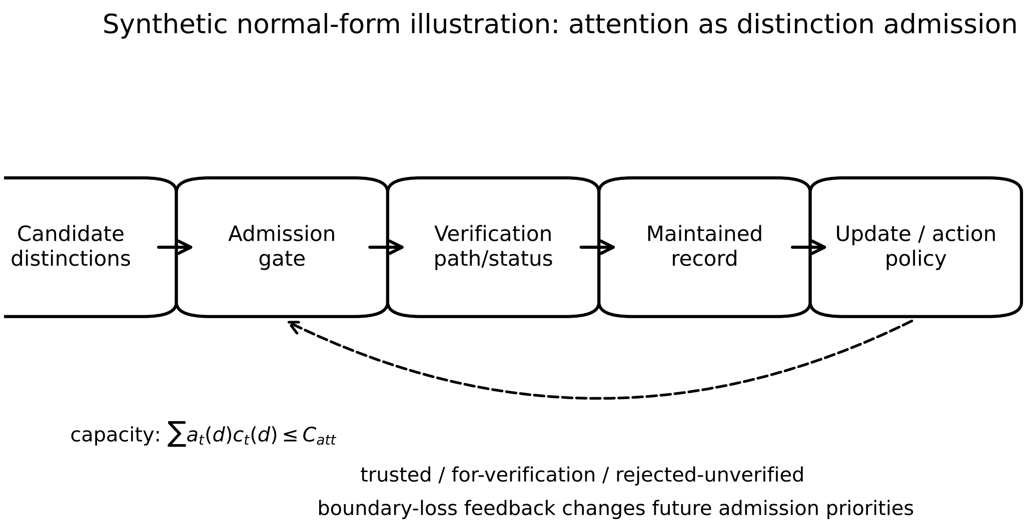

Figure 1: Synthetic normal-form illustration, not empirical evidence. Candidate distinctions pass through a finite admission gate before they can become verification-addressable records and condition downstream update or action. Boundary-loss feedback changes future admission priorities.

3.1 Maintained attention versus transient detection

A camera sensor may detect a photon without storing a stable record. A person may register a sound without using it. A model may compute an internal score for a token without preserving the resulting distinction in a durable update state. M1 reserves the term *attention* for admission into an update-relevant channel. This keeps attention distinct from raw sensing, transient activation, salience, and post-hoc report.

3.2 Verification status classes

Verification need not be complete before a distinction counts as attended. M1 distinguishes three cases:

Class	Meaning
Attended-as-trusted	The distinction is admitted and treated as reliable enough for downstream update.
Attended-for-verification	The distinction is admitted in order to test reliability, cross-check, or resolve conflict.
Rejected-unverified	The distinction is detected or available but not admitted because the verification path is too costly, unavailable, or outside the update window.

This distinction will matter for later work on trust, scientific error correction, and epistemic pollution.

3.3 Internal, externalized, and collective attention

Attention may be internal to a biological or artificial system; it may be externalized into lists, alarms, calendars, dashboards, or agendas; and it may be collective, as when a group allocates shared verification and communication capacity to some issues and not others. The same formal gate can describe these cases, but the accounting boundary must be stated.

4 Attention allocation under finite capacity

Each candidate distinction has an admission cost

$$c_t(d) = c_t^{\text{enc}}(d) + c_t^{\text{verify}}(d) + c_t^{\text{maint}}(d) + c_t^{\text{opp}}(d), \tag{6}$$

including encoding, verification, maintenance, and opportunity cost. Unless otherwise stated, $c_t(d)$ is a scalarized cost in loss-equivalent or budget-equivalent units. A more detailed model can replace Eq. (8) with vector constraints,

$$\sum_{d \in \mathcal{D}_t} a_t(d) c_t^r(d) \le C_t^r, \qquad r \in \{\text{enc}, \text{verify}, \text{maint}, \text{latency}, \text{energy}\}. \tag{7}$$

The scalar attention capacity constraint is

$$\sum_{d \in \mathcal{D}_t} a_t(d) c_t(d) \le C_{\text{att}}(t). \tag{8}$$

Definition 4 (Boundary-efficient attention allocation)**.** An attention allocation is boundary-efficient over horizon k if, under the stated model class and capacity constraints, no feasible reallocation lowers expected boundary-maintenance loss plus scalarized admission cost over the audited horizon:

$$a_t^\star \text{ is boundary-efficient if } \quad \forall a_t \in \mathcal{A}(C_{\text{att}}),$$
$$\mathbb{E}[\ell_{\text{maint,t+k}} \mid a_t^\star] + \lambda C(a_t^\star) \le \mathbb{E}[\ell_{\text{maint,t+k}} \mid a_t] + \lambda C(a_t) + \epsilon.$$

Definition 5 (Attention allocation)**.** An attention allocation is an admission vector a_t satisfying Eq. (8). A boundary-efficient finite attention system should allocate admission toward distinctions with high causal boundary relevance or high exploration value relative to their cost, while respecting capacity and verification constraints.

A normal-form allocation problem is

$$a_t^* \in \arg\max_{a_t} \sum_{d \in \mathcal{D}_t} a_t(d) \left[V_t^{\text{att}}(d;k) + \eta_t I_t^{\text{gain}}(d) \right] \quad \text{s.t.} \quad \sum_d a_t(d) c_t(d) \le C_{\text{att}}(t), \tag{9}$$

with $a_t(d) \in \{0,1\}$ for hard admission or $a_t(d) \in [0,1]$ for soft admission. The hard case is knapsack-like; the soft case is a fractional allocation. M1 does not claim real systems solve Eq. (9) exactly. It uses the equation as an audit normal form: deviations can be interpreted as salience capture, suppression, verification failure, learned priority, exploration, or misestimated value.

4.1 Exploration and background scanning

A system that admits only currently high-value distinctions can become brittle. It may miss weak early signals, model-breaking anomalies, or low-salience distinctions whose value is unknown. M1 therefore separates exploitation from background scanning:

$$C_{\text{att}}(t) = C_{\text{exploit}}(t) + C_{\text{scan}}(t), \tag{10}$$

where C_{exploit} is spent on known high-value distinctions and C_{scan} is reserved for uncertain or model-informative distinctions. A simple information-gain term is

$$I_t^{\text{gain}}(d) = H(\Theta_t \mid M_t) - \mathbb{E}_{o \sim P(o \mid d, M_t)} \left[H(\Theta_t \mid M_t, d, o) \right], \tag{11}$$

where Θ_t denotes latent task or environment variables and the expectation is taken over possible verification outcomes, measurement noise, source reliability states, or latent-state interpretations associated with admitting d. Equation (9) then formalizes an exploration-exploitation tradeoff: admitting only high known value can reduce short-run loss while increasing long-run model rigidity.

4.2 Verification cost from model incompatibility

Verification cost is not arbitrary. In an FDS observer, verification requires finite records, comparison operations, cross-checks, memory turnover, and possibly external sources [3]. A normal-form verification cost is

$$c_t^{\text{verify}}(d) = c_0 + \kappa_{\text{KL}} D_{\text{KL}}\big(P_{\text{ext}}(\cdot \mid d) \| P_{\text{int}}(\cdot \mid M_t)\big) + \kappa_{\text{src}} C_{\text{source}}(d) + \kappa_{\text{cross}} C_{\text{crosscheck}}(d). \quad (12)$$

The KL-like term represents incompatibility between the candidate signal and the internal model; the source and cross-check terms represent provenance and corroboration cost. The KL term is a normal-form contributor, not a universal law. In empirical applications it may be replaced by a bounded divergence, a likelihood-ratio conflict score, a source-discrepancy measure, or another model-discrepancy proxy. Surprising signals can be cheap to verify when the measurement channel is reliable, the source is trusted, or the relevant test is simple. Model incompatibility is only one contributor to verification cost: some surprising signals are cheap to verify, and some familiar signals are expensive to audit.

Proposition 1 (Attention-salience dissociation)**.** *Salience is not sufficient for attention. Under finite capacity, a salient distinction can be rejected if its expected boundary relevance is low or if its verification and maintenance cost is high.*

Proof sketch. Consider two candidate distinctions d_i and d_j when the capacity constraint permits only one full hard admission. Suppose raw salience satisfies $s(d_i) > s(d_j)$ but causal admission value satisfies $V_t^{\text{att}}(d_j; k) > V_t^{\text{att}}(d_i; k)$ after costs are included. Any allocation maximizing admitted causal value chooses d_j over d_i. Thus high salience does not imply attention. Under a fractional or value-density greedy rule, the same dissociation appears when $V_t^{\text{att}}(d_j; k)/c_t(d_j) > V_t^{\text{att}}(d_i; k)/c_t(d_i)$. □

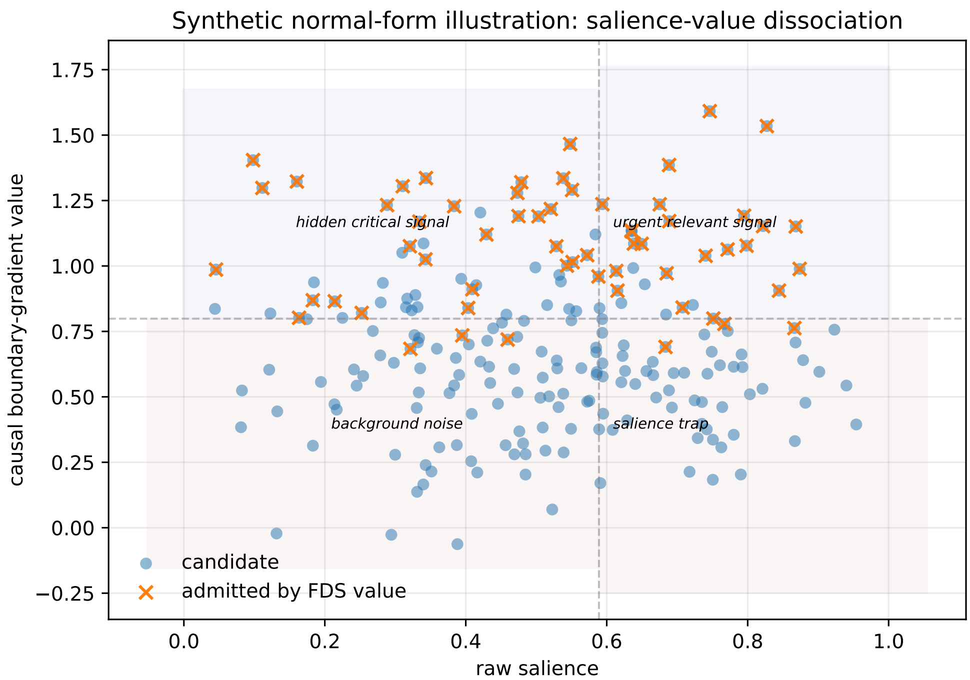

Figure 2: Synthetic normal-form illustration, not empirical evidence. Raw salience and causal boundary-gradient value can diverge. Quadrants distinguish salience traps from hidden critical signals. Under finite capacity, the admitted set can include lower-salience but higher-value distinctions.

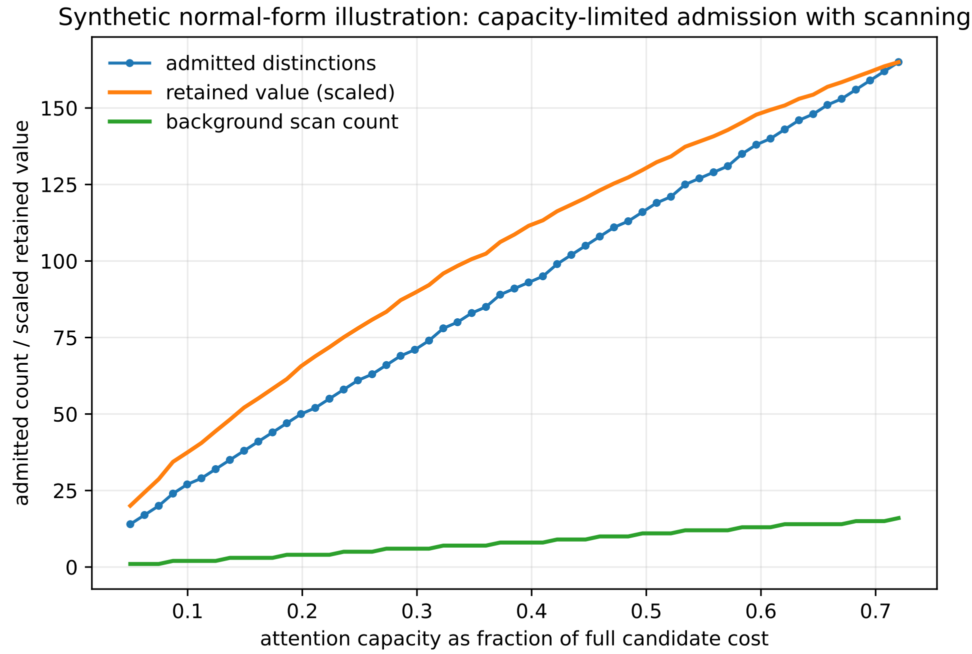

Figure 3: Synthetic normal-form illustration, not empirical evidence. As attention capacity increases, more distinctions can be admitted and more boundary-relevant value is retained. A reserved background-scanning budget admits some uncertain, high-information-gain distinctions that are not selected by exploitation alone.

5 Boundary relevance and causal attention value

FDS-M1 uses causal boundary relevance rather than mere correlation. Let $\mathrm{do}(\mathrm{admit}(d))$ denote an intervention that forces $a_t(d) = 1$ through the admission gate and permits downstream update and action policy to condition on the resulting maintained record. Let $\mathrm{do}(\neg\,\mathrm{admit}(d))$ force $a_t(d) = 0$ while leaving the rest of the gate, update, and policy mechanisms unchanged.

Definition 6 (Predictive attention relevance)**.** A distinction d has predictive attention relevance over horizon k if conditioning on it improves prediction of future boundary-maintenance loss:

$$V_t^{\mathrm{pred-att}}(d;k) = \mathbb{E}[\ell_{\mathrm{maint},t+k} \mid M_t] - \mathbb{E}[\ell_{\mathrm{maint},t+k} \mid M_t, d] - \lambda_t c_t(d). \tag{13}$$

Definition 7 (Causal boundary-gradient attention value)**.** A distinction d has causal attention value if admitting it changes expected future boundary-maintenance loss under the system's update and action channel:

$$V_t^{\mathrm{att}}(d;k) = \mathbb{E}[\ell_{\mathrm{maint},t+k} \mid M_t, \mathrm{do}(\neg\,\mathrm{admit}(d))] - \mathbb{E}[\ell_{\mathrm{maint},t+k} \mid M_t, \mathrm{do}(\mathrm{admit}(d))] - \lambda_t c_t(d). \tag{14}$$

The predictive version is an observational proxy. The causal version is preferred because a distinction that correlates with loss but cannot alter update, verification, action, or coordination is not yet attention-relevant in the operational sense.

Near collapse thresholds, average loss is insufficient. Let ℓ_c be a critical boundary-loss threshold. Write $A_d = \mathrm{admit}(d)$ and $\bar{A}_d = \neg\,\mathrm{admit}(d)$. Define

$$\Delta_d E\ell = \mathbb{E}[\ell_{\mathrm{maint},t+k} \mid M_t, \mathrm{do}(\bar{A}_d)] - \mathbb{E}[\ell_{\mathrm{maint},t+k} \mid M_t, \mathrm{do}(A_d)], \tag{15}$$

$$\Delta_d P_c = P(\ell_{\mathrm{maint},t+k} > \ell_c \mid M_t, \mathrm{do}(\bar{A}_d)) - P(\ell_{\mathrm{maint},t+k} > \ell_c \mid M_t, \mathrm{do}(A_d)). \tag{16}$$

Risk-weighted attention value is then

$$V_t^{\mathrm{risk-att}}(d;k) = \Delta_d E\ell + \alpha_t \Delta_d P_c - \lambda_t c_t(d). \tag{17}$$

The first term is expected loss reduction. The second is collapse-risk reduction. Both are positive when admission reduces risk.

Proposition 2 (Verification-limited attention)**.** *A finite system may fail to attend to a distinction not because it cannot detect the distinction, but because it cannot verify or maintain it within the update window.*

Proof sketch. In a hard-admission system, if $c_t(d) > C_{\text{att}}(t)$, no feasible allocation can admit d while satisfying Eq. (8). In a soft-admission system, d may receive a fractional weight, but it cannot become maintained attention if the maximum feasible weight $C_{\text{att}}(t)/c_t(d)$ falls below the minimum weight $a_{\min}$ required for downstream update. Thus detection does not entail maintained attention. $\square$

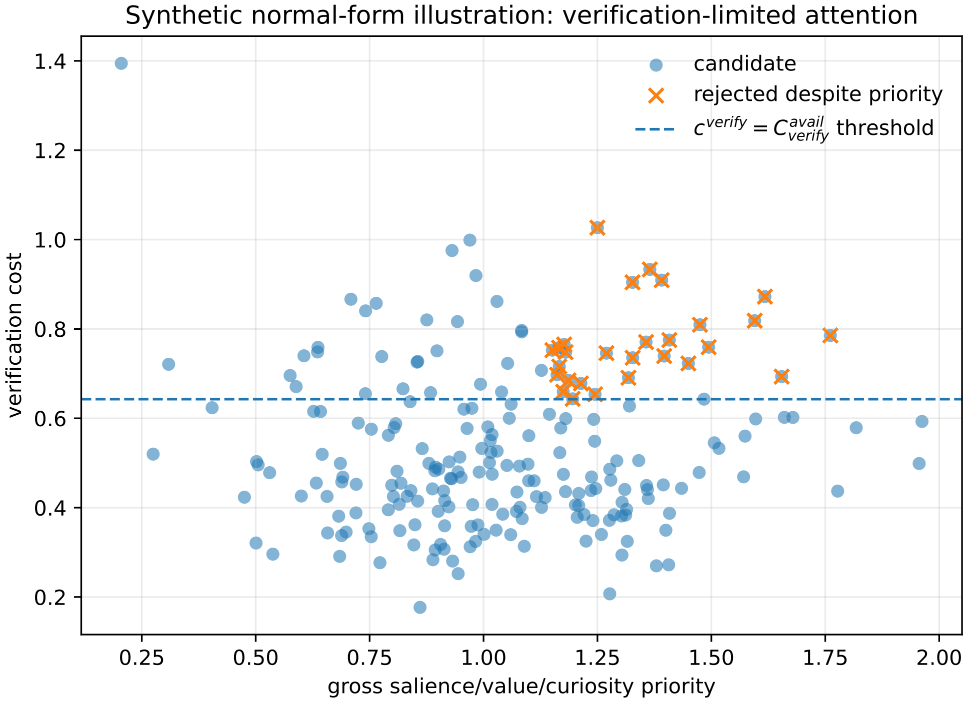

Figure 4: Synthetic normal-form illustration, not empirical evidence. Some distinctions have high gross priority but high verification cost. Candidates above the verification threshold cannot be admitted as maintained attention even when salience or gross priority is high.

5.1 Energy-time tradeoff and verification capacity

For short update windows, verification capacity can collapse even when raw detection capacity remains high. Fast attention therefore favors low-cost, high-priority admission rules and is more vulnerable to salience capture, stereotype-like compression, and risk-weighted shortcuts. A minimal bound is

$$C^{(\tau)}_{\text{verify}} \le R_{\text{verify}}\tau, \qquad N_{\text{checks}}(\tau) \le \frac{\tau}{\tau_{\text{check}}},$$

where τ_{check} is the minimum time per verification operation. When τ is small, verification capacity collapses and attention must rely on cheaper proxies. This connects M1 to the update-window constraints of O2 [5] and the O3 finite-memory operational channel [6]: rapid attention in short windows is not a cognitive flaw but a budget collapse.

6 Nonlinear attention gates, tunnel vision, and hysteresis

Attention gates often become nonlinear under load. Let $n_t(d)$ denote novelty, $r_t(d)$ threat or collapse-risk relevance, $I_t^{\mathrm{gain}}(d)$ information gain, and b_t a threshold. Define a gate score

$$V_t^{\mathrm{gate}}(d;k) = w_1 V_t^{\mathrm{att}}(d;k) + w_2 V_t^{\mathrm{risk-att}}(d;k) + w_3 I_t^{\mathrm{gain}}(d). \tag{18}$$

A normal-form gate is

$$a_t(d) = \sigma\left(\beta_t\left[V_t^{\mathrm{gate}}(d;k) + \alpha n_t(d) + \theta r_t(d) - \gamma c_t(d) - w_4 P_t^{\mathrm{false}}(d) C_t^{\mathrm{repair}}(d) - b_t\right]\right), \tag{19}$$

where $P_t^{\mathrm{false}}(d)$ is the probability that an admitted distinction is later judged false, misleading, or non-actionable, and $C_t^{\mathrm{repair}}(d)$ is the cost of repairing, retracting, or recovering from a false admission. False admission can be more expensive than non-admission because the system may need to maintain, propagate, retract, or repair an erroneous distinction. In social or institutional systems this repair cost may grow superlinearly with propagation depth. The gate steepness is

$$\beta_t = \beta_0 + \beta_\Delta[\Delta_{\mathrm{sem}}(t)]_+. \tag{20}$$

As semantic or attention deficit rises, the gate steepens. This produces tunnel vision: fewer distinctions are admitted, and admitted distinctions cluster around high immediate value, threat, or novelty. Long-horizon structure can be excluded even when it is important. This is the attentional analogue of capacity overflow and effective stochasticity in finite representation systems [4].

Proposition 3 (Load-induced attention narrowing)**.** *If gate steepness β_t increases with semantic deficit, then the admission probability becomes more threshold-like and the admitted set narrows around high-score distinctions.*

Proof sketch. For logistic $\sigma(\beta x)$, the derivative at the threshold is $\beta/4$ and the transition width scales as $1/\beta$. Increasing β_t therefore decreases the range of intermediate admission probabilities and produces a sharper split between admitted and rejected distinctions. □

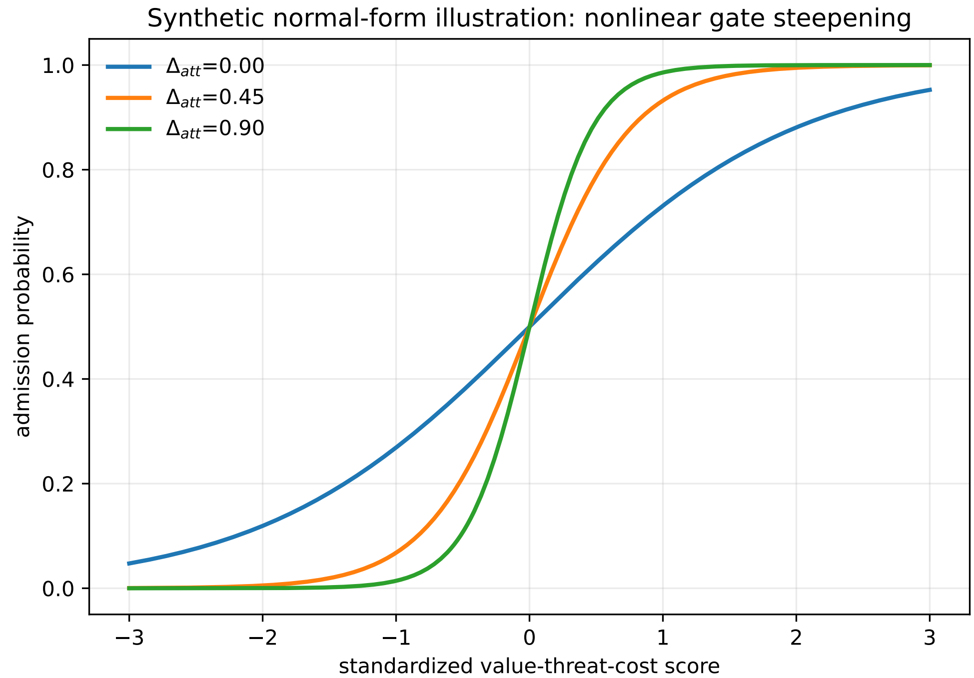

Figure 5: Synthetic normal-form illustration, not empirical evidence. As semantic or attention deficit increases, the admission gate steepens. Intermediate distinctions are excluded and the system enters a tunnel-vision regime.

6.1 Recovery delay and attention hysteresis

Tunnel vision can be useful when immediate boundary threat dominates; it becomes harmful when it blocks the distinctions needed for recovery. FDS-M1 treats recovery as an exit problem. The system can broaden attention by increasing C_{att}, reducing verification load, pruning obsolete candidates, compressing repeated distinctions, externalizing records, or relaxing the task.

Recovery need not be immediate. Crisis load can change priors, routines, verification checklists, or institutional thresholds. A minimal hysteresis model is

$$\beta_{t+1} = \beta_t + \rho\big(\beta_{\text{target}}(\Delta_{\text{sem}}(t)) + h_t - \beta_t\big), \qquad h_{t+1} = \chi h_t + \zeta \mathbf{1}_{\Delta_{\text{sem}}(t) > \Delta_c}, \tag{21}$$

where h_t is a gate-locking residue and $0 < \chi < 1$. Even after Δ_{sem} falls, h_t can keep the gate narrow for a recovery interval.

Proposition 4 (Exit-induced broadening with possible hysteresis)**.** *Under Eq. (19), an intervention that lowers Δ_{sem}, lowers $c_t(d)$, or increases C_{att} weakly broadens intermediate admission unless offset by threshold increase or hysteretic lock-in. Under Eq. (21), broadening can lag behind load reduction.*

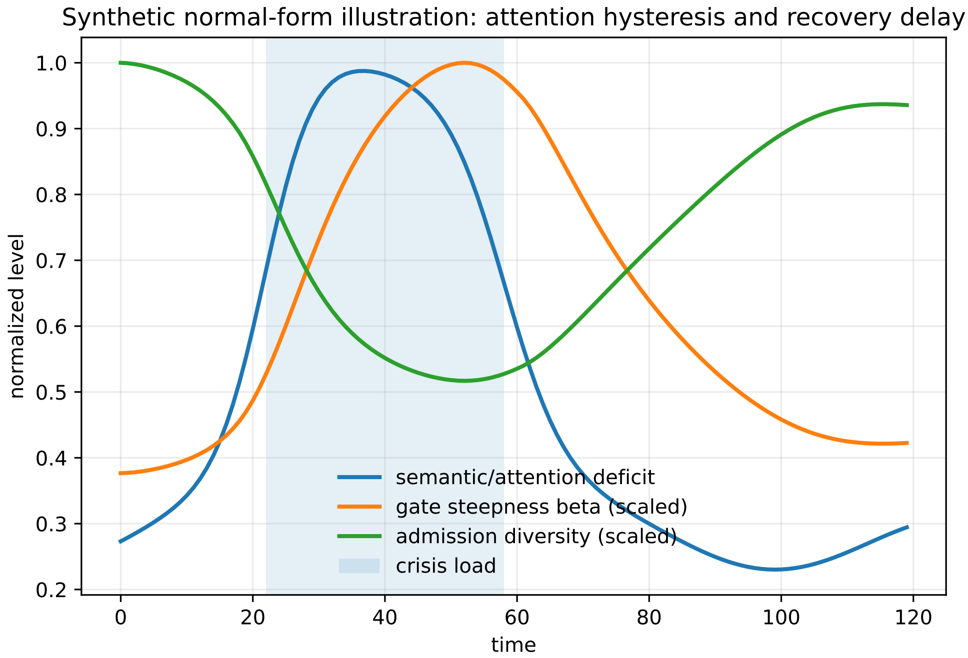

Figure 6: Synthetic normal-form illustration, not empirical evidence. During crisis load, gate steepness rises and admission diversity falls. After load reduction, attention can recover with a delay because gate thresholds, verification routines, or threat priors relax slowly.

7 Attention failure modes

FDS-M1 classifies attention failures as admission failures under finite capacity. This does not replace clinical, cognitive, or neural taxonomies; it provides an operational bridge.

Definition 8 (Attention overload)**.** Attention overload occurs when candidate distinction demand exceeds admission, verification, or maintenance capacity such that task-relevant distinctions are excluded.

Definition 9 (Distraction)**.** Distraction occurs when low boundary-relevance distinctions are admitted at the expense of higher boundary-relevance distinctions.

Definition 10 (Salience capture)**.** Salience capture occurs when raw salience dominates admission despite low causal boundary value.

Definition 11 (Suppression)**.** Suppression occurs when a candidate distinction is blocked despite positive boundary-gradient value, often because it threatens another maintained goal, identity, institution, or proxy objective.

Definition 12 (Critical distinction exclusion)**.** Critical distinction exclusion occurs when a distinction needed to prevent boundary failure is not admitted before the relevant update window closes.

Definition 13 (False admission)**.** False admission occurs when a low-quality, adversarial, unverifiable, or obsolete distinction enters the update channel and consumes capacity or corrupts downstream inference.

Definition 14 (Attention recovery lag)**.** Attention recovery lag occurs when a previously narrowed gate remains narrow after external load or threat has fallen, because verification routines, threat priors, or institutional thresholds relax more slowly than the triggering deficit.

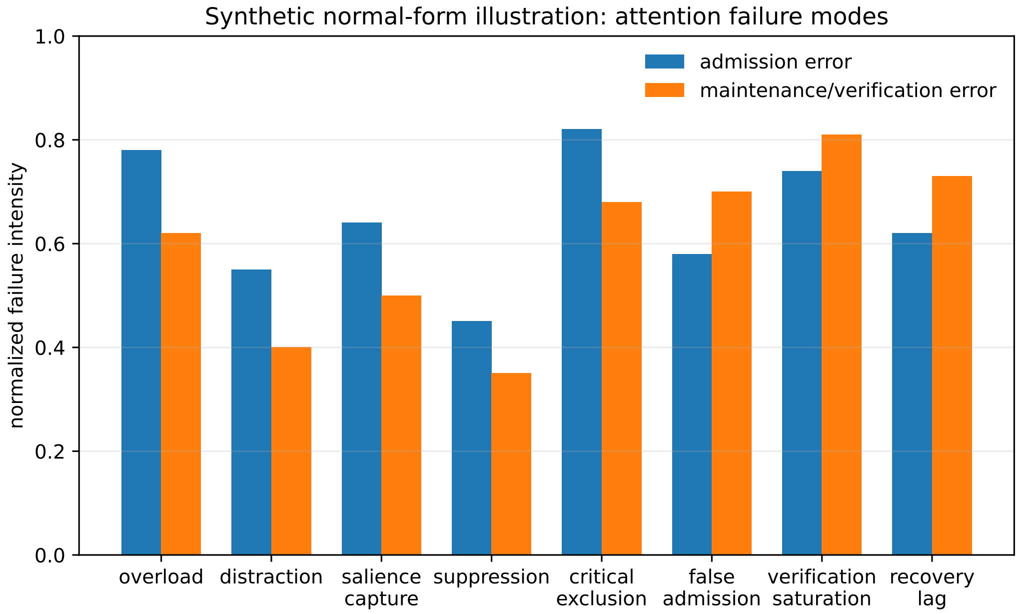

Figure 7: Synthetic normal-form illustration, not empirical evidence. Attention failures can be represented as admission errors, maintenance/verification errors, and recovery-lag errors under finite capacity.

8 Attention across domains

8.1 Biological and cognitive attention

In biological and cognitive systems, attention includes orienting, selection, working-memory gating, task-set maintenance, and action preparation. FDS-M1 does not choose among neural mechanisms. It supplies a shared accounting: a candidate distinction is attended only when it enters a maintained channel that can condition update or action. This connects limited-capacity attention to boundary relevance without claiming that all attention is consciously reportable.

8.2 Machine attention and transformer mechanisms

Attention weights in machine learning systems are not automatically FDS attention. Transformer attention is internal parallel routing among representations. FDS attention is budgeted admission into durable, update-relevant channels under a stated boundary, loss, and verification regime. A bare inference call may implement useful routing while lacking durable self-updating boundary state. A tool-using system with writable memory may satisfy stronger coupled-attention criteria if admitted distinctions affect future update and verification.

Dimension	Transformer attention	FDS attention
Basic role	token-to-token routing or weighting	distinction admission into update
Budget	compute/context constrained	encoding, verification, maintenance, latency, and boundary budget
Boundary	often implicit or absent	must be specified
Verification	usually not intrinsic	verification path/status is central
Agency attribution	not implied by weights	belongs to a coupled system if admitted records affect future loss

One FDS-compatible interpretation is that hallucination risk increases when many signals are routed or retrieved without a sufficient verification channel tied to a boundary-maintenance task. This is not a complete theory of hallucination; it is an admission-and-verification failure mode.

8.3 Organizational attention

Organizations attend through agendas, dashboards, meetings, metrics, issue queues, audits, and formal reporting channels. An issue may be known locally yet fail organizational attention if it does not enter a maintained decision channel. Organizational tunnel vision occurs when crisis, bureaucracy, or incentive misalignment steepens admission thresholds so that only short-horizon or proxy-relevant distinctions enter decision-making.

8.4 Civilization-scale attention

Civilization-scale attention is shared admission under finite communication, archival, scientific, legal, journalistic, and institutional verification capacity. Claims become collectively attended when they enter public records, standards, policy processes, scientific literature, or other externalized memory. Epistemic pollution is a saturation attack on finite verification bandwidth by low-cost, high-salience, low-reliability distinctions.

Let $R^{(\tau)}_{\mathrm{verify}}(\epsilon; t)$ be the verification demand imposed by the shared distinction environment and $C^{\mathrm{avail}}_{\mathrm{verify}}(t)$ the available civilization-scale verification capacity. Define

$$Z_{\mathrm{att/poll}}(t) = \frac{R^{(\tau)}_{\mathrm{verify}}(\epsilon; t)}{C^{\mathrm{avail}}_{\mathrm{verify}}(t)}. \tag{22}$$

When $Z_{\mathrm{att/poll}} > 1$, the system cannot verify all candidate claims. It must ignore, trust blindly, outsource verification, false-compress, or narrow the collective admission gate.

Proposition 5 (Collective attention as shared admission). *Collective attention is shared distinction admission under finite communication, verification, and externalized memory capacity. It fails when the shared gate admits low-quality distinctions or excludes boundary-critical distinctions at scale.*

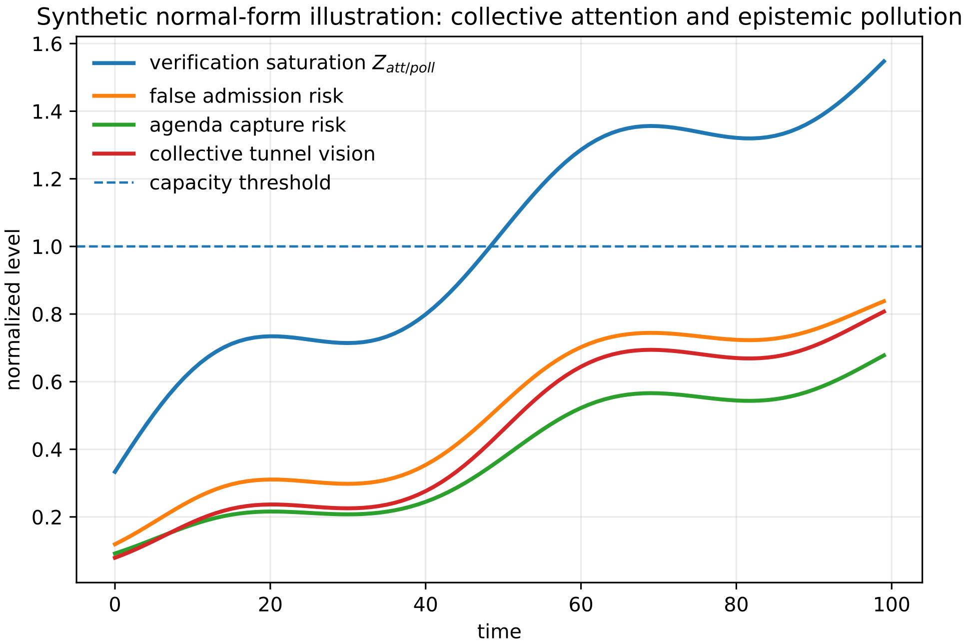

Figure 8: Synthetic normal-form illustration, not empirical evidence. Collective attention fails when verification demand saturates available verification capacity, increasing false admission, agenda-capture risk, and collective tunnel vision.

8.5 Domain mapping examples

Table 2: Illustrative domain mappings. Each row requires its own bridge assumptions and audit protocol.

Domain	Candidate distinction	Gate	Capacity	Failure
Cell	chemical gradient	receptor/regulatory gate	metabolic/regulatory bandwidth	missed nutrient or toxin
Human operator	alarm/anomaly	task set and working memory	attention and verification time	tunnel vision
LLM agent	tool output or memory record	context/memory admission	tokens and verification	hallucinated continuation
Organization	issue or warning	agenda/dashboard/audit	meeting and audit bandwidth	ignored risk
Civilization	claim or discovery	science/media/legal archive	verification institutions	epistemic pollution

9 Normal-form model and reproducibility

The accompanying code implements deterministic synthetic normal-form illustrations. The model is not empirical evidence. It is a consistency and visualization device that maps the definitions to simple state variables.

The model uses a candidate distinction stream with salience, novelty, threat, long-horizon relevance, source uncertainty, model divergence, verification cost, maintenance cost, encoding cost, causal boundary value, collapse-risk reduction, and information gain. It implements greedy capacity-limited admission, a reserved background-scanning budget, salience-value dissociation, capacity sweeps, verification-limited rejection, nonlinear deficit-dependent sigmoid gates, attention failure-mode intensities, hysteretic attention recovery, and collective attention saturation.

The random seed is fixed in `code/generate_results.py`. CSV outputs are stored in `data/`; figure pairs are stored in `figures/`.

Table 3: Normal-form variable map. All entries are illustrative, not fitted empirical quantities.

Simulation variable	Paper definition	Interpretation
salience	raw candidate prominence	detection strength, not attention itself
causal value	Eq. (14)	expected boundary-loss reduction minus admission cost
information gain	Eq. (11)	exploratory value of model-updating distinctions
verification cost	Eq. (12)	model incompatibility plus source and cross-check burden
semantic deficit	Eq. (20)	task demand pressure that steepens the gate
admission probability	Eq. (19)	soft gate output
recovery lag	Eq. (21)	delayed relaxation of attention narrowing
verification saturation	Eq. (22)	shared demand/capacity ratio

10 Protocols and tests

Protocol 1 (Admission audit)**.** Pre-register candidate distinctions, raw salience, costs, predicted boundary relevance, information gain, and admission outcomes. Test whether admitted distinctions are better predicted by causal boundary value and exploration value than by salience alone.

Protocol 2 (Salience-value dissociation test)**.** Construct cases where salience and boundary relevance diverge. M1 predicts that boundary-efficient attention should sometimes admit lower-salience but higher boundary-relevance distinctions.

Protocol 3 (Capacity load test)**.** Increase candidate distinction load or reduce admission capacity. Prediction: admission narrows, low-priority distinctions are excluded, and the admission gate becomes more thresholded.

Protocol 4 (Verification-burden test)**.** Increase verification cost while holding salience fixed. Prediction: unverifiable distinctions are rejected, delayed, externalized, or accepted blindly only under failure conditions.

Protocol 5 (Exploration-budget test)**.** Reduce or eliminate background scanning. Prediction: short-run performance may improve under stable conditions, but weak early signals, anomalies, and model-breaking distinctions are missed more often. Audit outputs include delayed anomaly detection rate, model-update surprise after missed scanning, and long-horizon boundary loss following short-horizon admission gains.

Protocol 6 (Tunnel-vision and hysteresis test)**.** Induce semantic deficit, crisis load, or threat pressure. Prediction: attention becomes steeply thresholded, high-threat distinctions dominate, and recovery of broad admission lags behind external load reduction when gate hysteresis is present.

Protocol 7 (Collective attention audit)**.** Measure how groups allocate shared attention under limited verification and communication capacity. Prediction: epistemic pollution increases false admission, agenda capture, false compression, or collective tunnel vision when $Z_{\mathrm{att/poll}} > 1$.

11 Relation to existing fields

Selective attention and limited capacity. Classical attention research already treats attention as selective and capacity-limited, including bottleneck, filter, and resource perspectives [7–9]. Recent reviews emphasize that selective processing spans multiple forms of attention, species, and levels of analysis [26]. FDS-M1 does not replace these accounts. It supplies a cross-domain finite-system interpretation: selection is admission into update.

Biased competition and priority maps. Neural and cognitive accounts often describe attention as competition among representations, biased by task goals and stimulus strength [10–12]. M1 is compatible with this: priority is interpreted as boundary-relevance weighted by cost and capacity.

Inattentional blindness and working memory. Inattentional blindness and working-memory limits show that detection and report are constrained by limited admission and maintenance [13–15]. Working-memory gating models similarly emphasize that information must pass controlled update gates to affect maintained state [16]. M1 formalizes this as distinction admission with maintenance and verification cost.

Rational inattention and information bottleneck. Rational inattention treats information acquisition as costly [17, 30]. The information bottleneck formalizes compression that preserves task-relevant information [18]. M1 uses both as nearby mathematical idioms, while tying admission to boundary-maintenance loss and update relevance.

Active inference and active sensing. Active inference and active sensing emphasize perception-action loops under uncertainty [19]. M1 is compatible with these approaches but emphasizes finite distinction capacity, verification burden, and the distinction between detection and admission.

Transformer attention. Transformer attention is a powerful computational selection and routing mechanism [20], but long context is not equivalent to robust admission: models can fail to use relevant information depending on its position in context [27]. Recent work on differential attention treats over-allocation to irrelevant context as an architectural failure mode [28]. M1 does not identify transformer attention with full FDS attention. The latter requires an accounting boundary, update relevance, and maintenance, verification, or downstream action effect. Recent analyses show that transformer mechanisms can acquire gating-like operations on working-memory tasks, but this does not by itself establish FDS attention unless the accounting boundary, durable update, and verification path are specified [29].

Organizational attention, situation awareness, and agenda-setting. Organizations and publics allocate attention through agenda formation, issue channels, and institutional filtering [21–23]. Situation awareness work analyzes how operators perceive, comprehend, and project task-relevant environmental elements [24]. Agenda-setting theory studies how media and public institutions shape issue salience at collective scales [25]. Credibility cues and social norms can improve truth discernment and reduce engagement with false content in social-media simulations, suggesting one empirical interface for collective verification capacity [31]. FDS-M1 interprets these as collective distinction admission under finite verification and communication capacity.

12 Limitations and falsification

M1 is intentionally limited. It does not establish a general empirical theory of attention by definition. It provides mappings that must survive operational audit. The framework is weakened or demoted under any of the following results:

1. attention-like selection under a specified mapping shows no finite capacity, admission, update gating, or priority constraint;
2. admission is always explained by raw salience and never by cost, verification, task relevance, exploration value, or boundary-gradient value;
3. increasing load produces no narrowing, thresholding, exclusion, or priority collapse in systems claimed to have finite attention;
4. distinctions can be detected, verified, maintained, updated, and acted upon without any capacity or resource tradeoff;
5. artificial attention weights alone satisfy FDS attention without an accounting boundary, downstream update relevance, or durable maintained record;
6. collective attention shows no relation to communication capacity, verification bandwidth, agenda channels, or externalized memory;
7. claimed attention hysteresis disappears under controlled recovery tests where load reduction should reveal persistent gate-locking.

13 Conclusion

Attention is the first active step in the agency-semantics chain. A distinction becomes operationally relevant only if it passes from availability into an update channel. M1 defines attention as capacity-limited distinction admission and separates it from raw salience, detection, and report. Its central chain is

$$\begin{aligned}\text{candidate distinction} &\to \text{admission gate} \to \text{verification path/status} \to \text{maintained record} \\ &\to \text{downstream update/action} \to \text{future boundary loss.}\end{aligned} \tag{23}$$

The paper introduces a finite-system vocabulary for salience-value dissociation, capacity-limited admission, background scanning, verification burden, tunnel vision, recovery hysteresis, machine-attention attribution, organizational attention, and collective attention under epistemic pollution.

M1 prepares later work. M2 can treat admitted distinctions as inputs to value and goal ranking. M3 can treat attention as the entry condition for actionable semantic quotients. M5 can analyze trust as delegated verification that reduces repeated admission cost. A2 can audit whether artificial systems externalize attention and verification to hosts. S2 can analyze epistemic pollution as saturation of civilization-scale attention and verification bandwidth.

Code and data availability

The deterministic synthetic normal-form code, figures, and CSV outputs are included in the replication package. Run `python code/generate_results.py` from the paper directory to regenerate all figures and data.

AI assistance disclosure

The author used AI assistance for drafting, editing, simulation scaffolding, and consistency checks. The author reviewed and selected the final claims, definitions, equations, and interpretations.

A M-series dependency map

Table 4: How M1 supports later M-series and civilization-layer papers.

Future paper	Interface supplied by M1
M2 Value/Goal	admitted distinctions become inputs to causal boundary-gradient ranking
M3 Meaning	attention supplies the entry condition for actionable semantic quotients
M5 Trust	trust reduces repeated verification cost and changes admission probability
A2 AI Alignment	artificial systems must state whether attention belongs to model, tool, memory, user, or coupled system
S2 Epistemic Pollution	pollution saturates shared verification bandwidth and distorts collective admission
G3 Science	scientific method functions as collective claim admission, verification, and error correction
G4 Civilization Memory	archives and standards stabilize externalized attention over time

B Notation summary

Table 5: Notation used in M1.

Symbol	Meaning
$\mathcal{D}_t$	candidate distinction stream at update window t
d	candidate distinction
$a_t(d)$	admission weight for distinction d
$c_t(d)$	scalarized encoding, verification, maintenance, and opportunity cost
$c_t^r(d)$	resource-specific vector cost
$C_{\text{att}}(t)$	attention admission capacity
$C_{\text{scan}}(t)$	reserved background-scanning capacity
$C_{\text{verify}}(t)$	verification capacity
$V_t^{\text{pred-att}}$	predictive attention relevance
V_t^{att}	causal boundary-gradient attention value
$V_t^{\text{risk-att}}$	risk-weighted attention value near collapse thresholds
V_t^{gate}	combined gate score used by the nonlinear admission gate
$I_t^{\text{gain}}(d)$	information gain / exploratory value of candidate d
$\Delta_{\text{sem}}(t)$	semantic capacity deficit
ℓ_{maint}	boundary-maintenance loss
Σ_{phys}	physical entropy-production ledger, distinct from ℓ_{maint}
$Z_{\text{att/poll}}$	collective verification saturation ratio
h_t	gate-locking residue / attention hysteresis state
χ	hysteresis persistence / decay factor, $0 < \chi < 1$
ζ	gate-locking accumulation increment under severe deficit
Δ_c	semantic deficit threshold above which gate locking accumulates
β_t	attention gate steepness
β_{target}	deficit-dependent target gate steepness

References

[1] Y. Wu, *Active Finite Distinction Systems: A Formal Core for Boundary Maintenance under Finite Capacity*, Zenodo (2026), doi:10.5281/zenodo.20158923.

[2] Y. Wu, *The Agency-Semantics Spine of Distinction Theory: Attention, Value, Goal, Meaning, and Action under Finite Capacity*, Zenodo (2026), doi:10.5281/zenodo.20257939.

[3] Y. Wu, *Observer as a Finite Distinction Register*, Zenodo (2026), doi:10.5281/zenodo.20248792.

[4] Y. Wu, *Capacity Overflow and Effective Stochasticity*, Zenodo (2026), doi:10.5281/zenodo.20250367.

[5] Y. Wu, *Time as Irreversible Distinction Update*, Zenodo (2026), doi:10.5281/zenodo.20249369.

[6] Y. Wu, *Boundary Maintenance and the Second Law under Finite Memory*, Zenodo (2026), doi:10.5281/zenodo.20255129.

[7] D. E. Broadbent, *Perception and Communication* (Pergamon Press, 1958).

[8] A. M. Treisman, "Monitoring and storage of irrelevant messages in selective attention," *Journal of Verbal Learning and Verbal Behavior* **3**, 449–459 (1964).

[9] D. Kahneman, *Attention and Effort* (Prentice-Hall, 1973).

[10] R. Desimone and J. Duncan, "Neural mechanisms of selective visual attention," *Annual Review of Neuroscience* **18**, 193–222 (1995).

[11] M. Corbetta and G. L. Shulman, "Control of goal-directed and stimulus-driven attention in the brain," *Nature Reviews Neuroscience* **3**, 201–215 (2002).

[12] E. I. Knudsen, "Fundamental components of attention," *Annual Review of Neuroscience* **30**, 57–78 (2007).

[13] A. Mack and I. Rock, *Inattentional Blindness* (MIT Press, 1998).

[14] D. J. Simons and C. F. Chabris, "Gorillas in our midst: sustained inattentional blindness for dynamic events," *Perception* **28**, 1059–1074 (1999).

[15] N. Cowan, "The magical number 4 in short-term memory: a reconsideration of mental storage capacity," *Behavioral and Brain Sciences* **24**, 87–114 (2001).

[16] R. C. O'Reilly and M. J. Frank, "Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia," *Neural Computation* **18**, 283–328 (2006).

[17] C. A. Sims, "Implications of rational inattention," *Journal of Monetary Economics* **50**, 665–690 (2003).

[18] N. Tishby, F. C. Pereira, and W. Bialek, "The information bottleneck method," in *Proceedings of the 37th Annual Allerton Conference on Communication, Control, and Computing* (1999).

[19] K. Friston, "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience* **11**, 127–138 (2010).

[20] A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems* **30** (2017).

[21] H. A. Simon, *Administrative Behavior* (Macmillan, 1947).

[22] R. M. Cyert and J. G. March, *A Behavioral Theory of the Firm* (Prentice-Hall, 1963).

[23] J. G. March and J. P. Olsen, "The uncertainty of the past: organizational learning under ambiguity," *European Journal of Political Research* **3**, 147–171 (1975).

[24] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors* **37**, 32–64 (1995).

[25] M. E. McCombs and D. L. Shaw, "The agenda-setting function of mass media," *Public Opinion Quarterly* **36**, 176–187 (1972).

[26] R. Xia, X. Chen, T. A. Engel, and T. Moore, "Common and distinct neural mechanisms of attention," *Trends in Cognitive Sciences* **28**, 554–567 (2024), doi:10.1016/j.tics.2024.01.005.

[27] N. F. Liu, K. Lin, J. Hewitt, A. Paranjape, M. Bevilacqua, F. Petroni, and P. Liang, "Lost in the Middle: How Language Models Use Long Contexts," *Transactions of the Association for Computational Linguistics* **12** (2024), doi:10.1162/tacl_a_00638.

[28] T. Ye, L. Dong, Y. Xia, Y. Sun, Y. Zhu, G. Huang, and F. Wei, "Differential Transformer," arXiv:2410.05258 (2024).

[29] A. Traylor, J. Merullo, M. J. Frank, and E. Pavlick, "Transformer Mechanisms Mimic Frontostriatal Gating Operations When Trained on Human Working Memory Tasks," arXiv:2402.08211 (2024), doi:10.48550/arXiv.2402.08211.

[30] A. Caplin, D. Martin, and P. Marx, "Modeling Machine Learning: A Cognitive Economic Approach," *Journal of Economic Theory* **224**, 105970 (2025), doi:10.1016/j.jet.2025.105970.

[31] T. Prike, L. H. Butler, and U. K. H. Ecker, "Source-credibility information and social norms improve truth discernment and reduce engagement with misinformation online," *Scientific Reports* **14**, 6900 (2024), doi:10.1038/s41598-024-57560-7.