Towards a Unified Operator Framework

First Author[1], Second Author[2]

[1] First author's department, institute, city, country.
[2] Second author's department, institute, city, country.

*corresponding author: email@email.com

Keywords: keyword1, keyword2, keyword3, keyword4, keyword5.

1 Introduction

Across decades of research, formal theories of human memory have proliferated into distinct and often incompatible frameworks. Localist serial-order models emphasize discrete positional codes (e.g., Burgess & Hitch, 1999; Henson, 1998), temporal context models describe memory as driven by a drifting contextual state (Howard & Kahana, 2002; Sederberg et al., 2008), and probabilistic accounts such as REM formalize recognition as stochastic sampling over noisy feature traces (Shiffrin & Steyvers, 1997). Each model family captures key empirical regularities—recency gradients, temporal clustering, distinctiveness effects, or probabilistic recognition—but they diverge in how they represent the basic building blocks of cognition: *content* and *context*.

This division has produced a persistent structural fragmentation in the theoretical landscape of memory. Models differ not only in their mechanisms but in the very geometry of their representational spaces: localist models assume symbolic sets of discrete items; connectionist and oscillator-based models operate on distributed feature vectors; and probabilistic frameworks treat both items and contexts as random variables sampled from latent feature distributions. Consequently, comparisons among models often confound differences in representational ontology with differences in process or mechanism (cf. Shiffrin, 2003; Nosofsky & Kantner, 2006; Busemeyer & Townsend, 1993). The field lacks a common mathematical substrate that can express symbolic, vectorial, and probabilistic representations within a single formal language.

The present work develops such a substrate. Building on prior integrative efforts that linked associative, sampling, and dynamic models under shared computational principles (e.g., Dubé & Malmberg, 2025; Farrell, 2021; Busemeyer & Townsend, 1993), we propose a general *operator framework* that unifies item–context binding across memory theories. Rather than adopting a specific algorithmic mechanism, we begin from first principles: (1) that memory encodes relations between elements of a content domain and a context domain, and (2) that these domains can each be represented as measurable vector spaces endowed with similarity metrics and probability measures. Within this framework, associative matrices, temporal drift dynamics, and probabilistic retrieval all emerge as constrained forms of a single operator family.

This unified operator approach is not a new model of memory, but a higher-level formalization that reveals structural equivalence among existing ones. It clarifies which empirical differences among theories reflect genuinely distinct assumptions about learning and retrieval, and which arise from representational parameterizations of the same underlying operator. In doing so,

Article version: 1st November 2025 h14:58 CET

it provides the foundation for expressing localist, connectionist, and probabilistic accounts of memory within a single, continuous mathematical system.

The following section introduces the general form of the Unified Operator Model and its measurable-space foundations.

Beyond its specific formalization, the significance of this work lies in establishing a representational common ground for cognitive theory. Rather than proposing another process-level model, the unified operator framework defines the structural conditions that make any model a *memory model*: a measurable relation between content and context, realized as a linear operator on their respective representational spaces. This provides a meta-theoretical language in which localist, connectionist, and probabilistic accounts can be compared not by their surface equations but by their underlying geometry. In this sense, the framework does not compete with existing models but situates them within a single operator family, clarifying which differences are substantive and which are parametrically equivalent. By treating item–context binding as an operator rather than a mechanism, the approach also bridges levels of analysis—linking symbolic, vectorial, and neural formulations through shared mathematical structure. This unification enables the derivation of novel predictions as continuous transformations within the operator space, offering a foundation for future theoretical synthesis across memory, decision, and learning domains.

1.1 Relation to Prior Operator and Vector-Space Frameworks

The present framework extends and generalizes several existing operator- and vector-space-based approaches to cognitive modeling.

Operator-theoretic approaches. Busemeyer and Townsend (1993) first introduced the notion of representing cognitive dynamics as linear operators acting on state vectors, thereby expressing sequential dependencies and decision processes within an algebraic framework. Their formulation, however, assumed a fixed representational space, such that the operator encoded temporal evolution within a single domain. In contrast, the Unified Operator Framework defines a mapping between *two* measurable domains—content and context—each represented by separable Hilbert spaces (H_F, H_Ψ). This shift elevates the operator concept from a process-level description to a representational meta-framework: associative matrices, temporal-drift dynamics, and probabilistic retrieval mechanisms all emerge as constrained instantiations of the same operator family $W = \sum_t \gamma_t f_t \otimes \psi_t$. Thus, whereas prior operator models formalized *within-space* transformations, the present approach captures *cross-space* binding as the fundamental structure underlying memory representations.

Vector-space and embedding approaches. The Unified Operator Framework also generalizes vector-based models such as BEAGLE (Jones & Mewhort, 2007), Holographic Reduced Representation (Plate, 1995), and semantic embedding models (Mikolov, Chen, Corrado, & Dean, 2013). These models represent associations as distributed bindings between continuous feature vectors, implemented through algebraic operations such as tensor products or circular convolutions. While these methods capture graded similarity and distributed representation, they typically fix the geometry of the representational space *a priori*. In the proposed operator framework, both item and context domains are measurable and can be continuously deformed under different encoding regimes. Binding is formalized not as a fixed algebraic product but as an operator integral over the joint measurable space $\Omega_F \times \Omega_\Psi$, allowing symbolic, vectorial, and probabilistic representations to coexist within a single formal system.

Summary. In sum, the Unified Operator Framework subsumes both linear-operator models and distributed vector frameworks as special cases defined by constraints on the representational measure, similarity metric, and operator domain. This unification clarifies which apparent theoretical differences among memory models reflect substantive psychological assumptions versus parameterizations of a common underlying operator structure.

1.2 Content and Context as Measurable Vector Spaces

Let Ω_F and Ω_Ψ denote the domains of representational *content* (items, percepts, or features) and *context* (temporal, spatial, or cognitive states). Each domain is treated as a measurable space,

$$(\Omega_F, \mathcal{F}_F, P_F), \qquad (\Omega_\Psi, \mathcal{F}_\Psi, P_\Psi),$$

where $\mathcal{F}_F$ and $\mathcal{F}_\Psi$ are σ–algebras of measurable subsets and P_F, P_Ψ define distributions over possible representational states. Mappings

$$f : \Omega_F \to \mathcal{H}_F, \qquad \psi : \Omega_\Psi \to \mathcal{H}_\Psi$$

embed these measurable elements into separable Hilbert spaces $\mathcal{H}_F$ and $\mathcal{H}_\Psi$, each equipped with an inner product $\langle \cdot, \cdot \rangle$. This embedding allows the same formal machinery to describe discrete localist codes (via Dirac measures), distributed feature vectors (via continuous embeddings), and probabilistic memory traces (via random measures).

1.3 The Memory Operator

Within this framework, a memory trace is defined as a linear operator on the tensor product space $\mathcal{H}_F \otimes \mathcal{H}_\Psi$, capturing the association between content and context:

$$W = \int_{\Omega_F \times \Omega_\Psi} f(x) \otimes \psi(y)\, d\mu(x, y),$$

where μ is an encoding measure over item–context pairs. In the discrete limit, this integral reduces to the summation

$$W = \sum_{t=1}^{T} \gamma_t f_t \otimes \psi_t,$$

recovering the common operator form underlying TCM, OSCAR, CRU, and Farrell's hierarchical model. Retrieval follows as a bilinear projection,

$$a(i|\text{cue}) = f_i^\top W\, \psi(\text{cue}),$$

in which activation reflects the alignment between the current context and the stored bindings. This operator equation serves as the foundation for unifying localist, connectionist, and probabilistic accounts of memory within a single mathematical framework.

1.4 Mapping Extant Models to the Operator Framework

The unifying power of the operator framework is best demonstrated by showing how prominent, and often competing, models of short-term and serial memory emerge as specific instantiations of the general operator equation. Recall the general form of the memory operator W and the retrieval activation $a(i|\text{cue})$:

$$W = \sum_{t=1}^{T} \gamma_t f_t \otimes \psi_t, \qquad a(i|\text{cue}) = \langle f_i, W\psi(\text{cue}) \rangle$$

The fundamental differences between model families lie not in this core associative structure, but in their definitions of the item representation f_t, the context representation ψ_t, and the encoding weight γ_t. We categorize these instantiations into three primary classes based on the nature of the context vector ψ_t.

Mechanics of the Memory Operator. The operator W acts as a linear transformation that maps a contextual state $\psi \in H_\Psi$ to a predicted content representation $\hat{f} \in H_F$:

$$\hat{f} = W\,\psi = \sum_t \gamma_t f_t \langle \psi_t, \psi \rangle.$$

Each stored pair (f_t, ψ_t) contributes a rank-one transformation $f_t \otimes \psi_t$, weighted by γ_t. The inner product $\langle \psi_t, \psi \rangle$ computes contextual similarity, while the weighted sum over t reconstructs a content vector as the superposition of associated features. Hence, retrieval corresponds to a projection of the cue context onto the subspace spanned by all stored contexts $\{\psi_t\}$, followed by re-expression in the content basis $\{f_t\}$.

Adjoint and reverse mapping. Because W is a bounded linear operator between Hilbert spaces, it possesses an adjoint $W^\top : H_F \to H_\Psi$ defined by

$$W^\top f = \sum_t \gamma_t \psi_t \langle f_t, f \rangle.$$

Psychologically, $W^\top$ performs the inverse process: it infers or reconstructs a contextual representation from a content vector. Thus, W and $W^\top$ form a dual pair implementing the two canonical directions of associative memory:

$$\psi \xrightarrow{W} f \quad \text{(context}\to\text{content retrieval)}, \qquad f \xrightarrow{W^\top} \psi \quad \text{(content}\to\text{context reinstatement)}.$$

A model is said to realize *bidirectional binding* when both mappings are active (e.g., recurrent or joint-embedding regimes), and *unidirectional binding* when only W is functionally used (e.g., product-space models).

Composition and superposition. Two operators can be composed to represent multi-stage processes. For example, applying W twice,

$$WW^\top : H_F \to H_F,$$

yields an autoassociator that projects a content vector onto the subspace of stored items. Similarly,

$$W^\top W : H_\Psi \to H_\Psi$$

projects a context vector onto the subspace of stored contexts. If $W^\top W \approx I_{H_\Psi}$, the operator is approximately unitary, meaning the mapping preserves contextual geometry and allows exact reconstruction—an idealized limit rarely achieved in psychological data.

Learning and updating. Encoding corresponds to additive composition of new rank-one operators:

$$W_{t+1} = W_t + \gamma_t f_t \otimes \psi_t.$$

Each new item–context pair updates the transformation, incrementally altering the mapping from contexts to contents. Forgetting or interference can then be expressed as post-multiplying W by a decay operator D_Ψ or pre-multiplying by D_F, both of which scale or distort the subspaces spanned by $\{f_t\}$ and $\{\psi_t\}$.

Geometric interpretation. Within this formalism, memory is a geometry-preserving correspondence between two vector manifolds. Each encoded trace defines a small linear map from a local patch of the context space to a local patch of the content space; the full memory operator is the superposition of these local maps. Retrieval is therefore not the selection of a discrete record but the evaluation of a continuous transformation whose output depends on the alignment between the probe context and the manifold of stored contexts. Different memory models can be understood as different constraints on the geometry and dimensionality of this mapping: orthogonal and static in positional models, metric-distorted in embedding models, and dynamically evolving in joint-embedding models.

2 Model Families as Operator Instantiations

The general operator, $W = \sum_{t=1}^{T} \gamma_t f_t \otimes \psi_t$, is not a new model itself, but a meta-framework. Its value is demonstrated by showing how major, seemingly disparate, model families emerge as specific instantiations of this equation. These families differ primarily in the mathematical constraints they place on the item (f_t) and context (ψ_t) vectors and, most critically, on the nature of their coupling. We identify three primary mathematical regimes that categorize the majority of extant STM models.

2.1 Regime 1: Product Space (Separable Item $\times$ Context)

In this regime, the item space $\mathcal{H}_F$ and context space $\mathcal{H}_\Psi$ are treated as separable. The retrieval function $a(i|\text{cue}) = f_i^\top W \, \psi(\text{cue})$ is fully bilinear, effectively calculating item-item similarity and context-context similarity independently and then combining them:

$$a(i|\text{cue}) = \sum_{t=1}^{T} \gamma_t \langle f_i, f_t \rangle \langle \psi(\text{cue}), \psi_t \rangle$$

This is the foundational assumption of positional-marking models. For example, in the Start-End Model (SEM; Henson, 1998) and the network model of Burgess & Hitch (1999), the context vector ψ_t is a discrete, item-independent positional marker (e.g., a localist vector $e_{p(t)}$ representing serial position $p(t)$). Similarly, oscillator-based models (e.g., OSCAR; Brown et al., 2000) instantiate ψ_t as a unique vector derived from the state of internal oscillators at time t.

The strength of this regime is its explicit coding for serial position, which naturally explains key positional error patterns such as transpositions, protrusions (Osth & Hurlstone, 2022), and grouping effects (Henson, 1998).

2.2 Regime 2: Embedding Space (Context-as-Metric)

This regime collapses the bilinear retrieval into what is effectively a unary operation. This is often achieved by treating ”context” not as a separate vector to be bound, but as an implicit distortion of the item-space metric itself.

The clearest example is the SIMPLE model (Brown et al., 2007). In SIMPLE, retrieval activation is purely a function of temporal distance in log-space: $a(i|\text{cue}) \propto \exp(-\lambda |\log t_i - \log t_{cue}|)$. In our framework, this corresponds to a model where the context vector ψ_t is simply the temporal position t, and the operator W is structured such that $\langle f_i, W\psi_{cue} \rangle$ computes a log-distance metric. This regime also includes exemplar-based recognition models (e.g., Nosofsky et al., 2011), where activation is an accumulated sum of similarity between a probe and memory traces, often without an explicit, independent context representation. These models excel at explaining scale-invariant recency and list-length effects (Brown et al., 2007).

2.3 Regime 3: Joint Embedding (Co-Evolving Item-Context)

In this third regime, the item and context spaces are inextricably linked and co-evolve over time. The context ψ_t is not a static marker but is itself a product of the item sequence.

The archetypal model here is the Temporal Context Model (TCM; Howard & Kahana, 2002). In TCM, the context vector ψ_{t+1} is a dynamic integration of the previous context and the just-presented item: $\psi_{t+1} = \rho\psi_t + \eta f_t$. This dynamic makes the operator $W = \sum f_t \otimes \psi_t$ a store of associations between items and the *compounded* context that preceded them. This mechanism is also central to the Context Retrieval and Updating model (CRU; Logan, 2021), where context is explicitly modeled as a mixture of previous item representations.

The critical empirical contribution of this regime is its ability to naturally explain the asymmetric contiguity effect (Howard & Kahana, 2002), where recall of one item tends to cue the *next* item in the list—a finding that is challenging for static product-space models.

3 Binding Regimes and Model Correspondence

The Unified Operator Framework (UOF) distinguishes among families of memory models based on how they instantiate *binding* between content and context. Each regime reflects specific assumptions about the independence or dependence between item representations (f_t) and contextual states (ψ_t). Formally, all regimes can be expressed within the general operator form

$$W = \sum_t \gamma_t f_t \otimes \psi_t,$$

where W maps from a context space H_Ψ to an item space H_F. By varying the generative assumptions of ψ_t, classical models of memory emerge as special cases or limiting forms of this operator equation.

3.1 Independence Assumptions Across Model Families

The following table summarizes how the three principal UOF regimes—and their extensions—map onto established model families, highlighting their binding assumptions and the mathematical conditions under which they are equivalent to the operator form.

3.2 Interpretation

The UOF unification clarifies that major memory models occupy distinct positions along a continuum of *context dependence*:

- **Context-independent binding** (Regime 1): items associate with pre-existing, item-independent context codes.
- **Implicit context or metric binding** (Regime 2): context collapses into a distortion of item–item similarity.
- **Item-dependent binding** (Regime 3): context dynamically evolves with the presented items.

Models that omit context altogether (e.g., REM, SAM) or incorporate recurrent context updates (e.g., SOB, connectionist frameworks) represent boundary conditions of the same operator family. The operator expression

$$a(i|\text{cue}) = \langle f_i,\, W\, \psi_{\text{cue}} \rangle$$

serves as a universal retrieval equation encompassing these regimes. Two models are operator-equivalent if their retrieval rule can be cast in this bilinear form for some parameterization of

Table 1: Binding regimes in the Unified Operator Framework and correspondence with established models.

UOF Regime	Binding Nature (Assumption)	Representative Models	Mathematical Equivalence within UOF
Regime 1: Product Space (Separable Item $\times$ Context)	Context is independent of items; binding links each item f_t to a pre-specified context tag ψ_t.	Henson's Start–End Model (1998); Burgess & Hitch (1999); OSCAR (Brown et al., 2000).	Retrieval $a(i\|\text{cue}) = \sum_t \gamma_t \langle f_i, f_t \rangle \langle \psi(\text{cue}), \psi_t \rangle$. Equivalent to associative-matrix form when ψ_t are orthogonal position markers.
Regime 2: Embedding Space (Context-as-Metric)	No explicit context variable; context effects arise through a metric deformation in item space.	SIMPLE (Brown, Neath, & Chater, 2007); exemplar/global-matching models (Nosofsky, 1986; 2011).	Can be expressed as a degenerate operator where $\psi_t = \psi(f_t)$. The bilinear form collapses to a unary similarity kernel $a(i\|\text{cue}) = \kappa(d(f_i, f_{\text{cue}}))$.
Regime 3: Joint Embedding (Co-evolving Item–Context)	Context depends on preceding items; $\psi_{t+1} = \rho\psi_t + \eta f_t$.	Temporal Context Model (Howard & Kahana, 2002); Context Retrieval and Updating (CRU; Logan, 2021).	Exactly matches the operator definition: $W = \sum_t f_t \otimes \psi_t$, retrieval $a_i = f_i^\top W \psi_{\text{cue}}$.
Extended Case A: Item–Item Association (No Context)	Binding confined to item space; context absent or static.	REM (Shiffrin & Steyvers, 1997); SAM (Raaijmakers & Shiffrin, 1980).	Equivalent to a self-associative operator $W = \sum_t f_t \otimes f_t$, representing direct feature–feature association.
Extended Case B: Recurrent or Feedback Context	Context evolves through retrieval or neural feedback: $\psi_t = g(\psi_{t-1}, f_{t-1})$.	Connectionist recurrent models; SOB (Farrell & Lewandowsky, 2002).	Representable via recursive operator updates $W_{t+1} = W_t + \Delta(f_t, \psi_t)$, approximating recurrent context dynamics.

W, f_t, and ψ_t. Under this criterion, positional and temporal-context models correspond to *exact* operator instantiations, while distinctiveness-based and global-matching models correspond to *limiting or degenerate* cases in which context collapses into the item space.

3.3 Theoretical Implication

The UOF thereby provides a principled mathematical substrate linking symbolic, vectorial, and probabilistic memory theories. What differ among models are not their fundamental associative structures but the *dependency assumptions* that generate their context vectors. This yields a continuous manifold of models parameterized by the function $\psi_t = g(\psi_{t-1}, f_{t-1})$, uniting the classic categories of context-independent, item-dependent, and item–item association within a single operator family.

4 Forgetting: A Unified Operator Perspective

Forgetting in memory research has long been partitioned into apparently distinct mechanisms— trace decay, interference, temporal distinctiveness loss, and contextual drift. Within the Unified Operator Framework, these mechanisms are revealed as alternative *placements of a loss operator* within the same compositional equation:

$$a(i|\text{cue}) = f_i^\top (D_f\, W\, D_\psi)\, \psi(\text{cue}), \tag{1}$$

where D_f and D_ψ are bounded linear decay or distortion operators acting respectively on the item and context spaces, and W is the item–context binding operator. All empirical forms of forgetting correspond to reductions in the overlap between the study-phase and test-phase representations; they differ only in the locus of this attenuation within the operator chain.

4.1 Historical Dimensions of Forgetting

Classical theories emphasized distinct loci of loss:

- **Trace-decay theories** (Atkinson & Shiffrin, 1968) posit magnitude reduction in W itself: $W_{t+1} = (1 - \lambda)W_t$.
- **Interference and cue-overload accounts** (Postman & Underwood, 1973) treat forgetting as overlap among stored pairs, i.e., $\langle f_t, \psi_t\rangle$ non-orthogonality.
- **Temporal distinctiveness models** such as SIMPLE (Brown, Neath, & Chater, 2007) interpret forgetting as compression of the temporal metric: $D_\psi = \exp[-\lambda|\log t_i - \log t_j|]$.
- **Context-drift models** (Howard & Kahana, 2002; Mensink & Raaijmakers, 1988) express forgetting as gradual decorrelation of context states, $\psi_{t+1} = \rho\psi_t + \eta f_t$, so that $\langle \psi_{\text{study}}, \psi_{\text{test}}\rangle$ decreases with lag.
- **Novelty-gated and attentional models** such as SOB (Farrell & Lewandowsky, 2002) assign decay to the encoding weights γ_t.
- **Connectionist and neural models** (Davelaar et al., 2005; Anderson et al., 2004) place loss on synaptic weights, equivalent to weight decay in W.

Although these frameworks differ terminologically, they all implement a common mathematical transformation: a contraction of representational similarity either in the content space (HF), the context space (H_Ψ), or the cross-space binding W. This equivalence exposes forgetting not as a process distinction but as a topological one—an operation that reduces the inner-product overlap in a Hilbert manifold.

4.2 A Multi-Dimensional Taxonomy of Forgetting

The operator view identifies three continuous dimensions along which models vary:

(i) **Spatial locus** — item, context, or operator. Determines whether loss manifests as feature fading, contextual drift, or synaptic degradation.

(ii) **Temporal regime** — instantaneous vs. gradual. Determines whether forgetting follows exponential decay ($e^{-\lambda t}$) or power-law compression (Wixted, 2004).

(iii) **Stochastic regime** — deterministic vs. probabilistic. In probabilistic accounts such as REM (Shiffrin & Steyvers, 1997) or EBRW (Nosofsky, 1991), decay is represented as variance inflation in the sampling distribution rather than deterministic contraction.

These axes generate a continuous *operator manifold* within which classical models occupy distinct regions: SIMPLE and temporal distinctiveness models correspond to metric distortion on the temporal axis; TCM and CRU occupy the context-drift plane; SOB and novelty-gated models occupy the encoding-weight dimension; and connectionist accounts trace the weight-decay diagonal. Hybrid placements—e.g., simultaneous metric compression and context drift—represent empirically unexplored regions of the manifold.

4.3 Theoretical Implications

Recasting forgetting as operator placement yields three implications:

(a) **Unified algebra:** All forms of forgetting can be represented as compositions of bounded linear maps within the operator algebra $\mathcal{M} = \{W, D_f, D_\psi\}$.

(b) **Parameter continuity:** Empirical forgetting functions that appear distinct (exponential vs. power-law) emerge as different parameterizations of the same operator family.

(c) **Predictive synthesis:** The placement formalism enables derivation of novel predictions by continuously shifting the decay locus, thus bridging verbal distinctions such as "decay vs. interference."

This operator generalization aligns with evidence that memory loss depends jointly on temporal discriminability, contextual reinstatement fidelity, and encoding strength (Neath & Brown, 2012; Sederberg et al., 2008). In this light, forgetting is not a failure of memory mechanisms but an expected geometric consequence of how similarity and measure evolve within the joint content–context space.

Proposition 1 (Operator Placement Equivalence)**.** *Consider the general memory operator*

$$W = \sum_{t=1}^{T} \gamma_t f_t \otimes \psi_t,$$

where $f_t \in H_F$ and $\psi_t \in H_\Psi$. Let forgetting be implemented by a linear transformation on one of three loci:

(i) context decay: $\psi_t' = D_\Psi \psi_t$,

(ii) encoding decay: $\gamma_t' = d_t \gamma_t$,

(iii) trace decay: $W' = D_W W$.

Then there exists a bounded linear operator $T : H_\Psi \to H_\Psi$ such that, up to a scalar normalization constant,

$$W' = WT,$$

where T encodes the effective locus of decay. Thus, decay, interference, and distinctiveness loss correspond to different parameterizations of the same composite transformation on the memory operator.

Sketch. Each form of decay multiplies one component of the tensor product $f_t \otimes \psi_t$ by a scalar or linear map. By linearity, this multiplication can be equivalently represented as right-multiplication of the composite operator by $T = \sum_t d_t\, \psi_t \psi_t^\top$. Hence all three forms of forgetting are isomorphic under reparameterization. □

5 From Global Matching to Recall: Conditions and Operator Transitions

Let the memory system be formalized on the product space $\mathcal{X} \times \mathcal{C}$, where $\mathcal{X}$ denotes the content (item) manifold and $\mathcal{C}$ denotes the contextual manifold. Each encoded trace is represented as a measure or vector

$$m_i = (x_i, c_i) \in \mathcal{X} \times \mathcal{C},$$

and the probe at test is given by

$$p = (x_p, \hat{c}_p),$$

where $\hat{c}_p$ is an estimate or reinstatement of the current context.

(i) Global Matching. Recognition is achieved by evaluating a global similarity operator

$$\mathcal{F}(p) = \sum_{i=1}^{N} K\big((x_p, \hat{c}_p), (x_i, c_i)\big),$$

where K is a similarity kernel separable in content and context dimensions,

$$K\big((x, \hat{c}), (x', c')\big) = K_X(x, x') \cdot K_C(\hat{c}, c').$$

If the contextual kernel K_C is sufficiently high-dimensional and sharply peaked, then only traces with matching or nearby c_i values contribute non-trivially to $\mathcal{F}(p)$. Formally, this is still a *global* operation: all traces are evaluated, but most have negligible weight. Recognition decisions are made by comparing $\mathcal{F}(p)$ to a criterion or by letting $\mathcal{F}(p)$ drive an evidence accumulator (e.g., EBRW).

(ii) Context-Gated Matching. A recall-like process emerges when the system first applies a *contextual projection operator*

$$P_{\hat{c}_p} : \mathcal{M}(\mathcal{X} \times \mathcal{C}) \to \mathcal{M}(\mathcal{X}), \qquad P_{\hat{c}_p}[m_i] = \delta(c_i - \hat{c}_p)\, m_i,$$

which restricts the trace set to those associated with the reinstated context $\hat{c}_p$. Similarity is then computed only within the gated subset:

$$\mathcal{R}(p) = \sum_{i:\, c_i = \hat{c}_p} K_X(x_p, x_i).$$

The identity of the maximal contributor $i^* = \arg\max_i K_X(x_p, x_i)$ can be used to *recover* additional latent information (e.g., list membership, serial position). This recovery step defines recall within the unified operator system.

Lemma 1 (Transition from Recognition to Recall)**.** *Let $\mathcal{F}$ and $\mathcal{R}$ be defined as above. A global-matching process becomes recall when the evaluation operator factorizes into a contextual projection followed by a content-based identification:*

$$\mathcal{R} = (\operatorname*{Id}_X \circ P_{\hat{c}_p})[\mathcal{F}],$$

such that the system outputs a recovered latent variable from the restricted subset rather than a scalar familiarity. Equivalently, recall is achieved when the context operator $P_{\hat{c}_p}$ changes order with the global sum, i.e.,

$$P_{\hat{c}_p}\Big(\sum_i K\Big) \neq \sum_i P_{\hat{c}_p}(K),$$

so that evaluation becomes local in $\mathcal{C}$ rather than global.

In summary, high-resolution context weighting (approach i) maintains global matching by embedding contextual information within the similarity kernel itself, whereas context gating (approach ii) introduces a projection operator that effectively selects a contextual subspace prior to matching, thereby instantiating recall.

Theorem 2 (Recognition–Recall Equivalence)**.** *Let the memory operator be defined as*

$$F(p) = \sum_{i=1}^{N} K_X(x_p, x_i) K_C(c_p, c_i),$$

where K_X and K_C are positive-definite kernels on the content and context spaces H_F and H_Ψ, respectively. Define the recall operator as the contextual projection

$$R(p) = \sum_{i:\, \|c_i - c_p\| < \varepsilon} K_X(x_p, x_i).$$

If the contextual kernel K_C is Gaussian with variance parameter σ^2,

$$K_C(c_p, c_i) = \exp\left[-\frac{\|c_p - c_i\|^2}{2\sigma^2}\right],$$

then in the limit as $\sigma \to 0$,

$$\lim_{\sigma \to 0} F(p; \sigma) = R(p).$$

Hence, recall emerges as the high-resolution limit of recognition within the same operator family.

Intuition. As the contextual kernel sharpens, all nonmatching contexts receive exponentially vanishing weight, so the global similarity sum becomes a projection onto the local context manifold. Formally, $K_C(\cdot, \cdot) \to \delta(\cdot - \cdot)$, turning the global integral into a restricted sum. Thus recognition and recall correspond to continuous operator transitions governed by the resolution of the contextual kernel. □

Theorem 3 (Dimensionality Threshold for Recall)**.** *Let $\{\psi_i\}_{i=1}^N \subset \mathbb{R}^{d_\Psi}$ be independent, unit-norm context vectors with pairwise cosine similarity $\rho = \mathbb{E}[\langle \psi_i, \psi_j \rangle]$, $i \neq j$. Let recall succeed when the contextual projection*

$$P_{\hat{c}} = \arg\max_i \langle \psi_i, \hat{c} \rangle$$

returns a unique match. Then the probability of unique recall satisfies

$$\Pr[\textit{unique recall}] \approx 1 - N\, e^{-d_\Psi (1-\rho^2)/2}.$$

Consequently, the critical dimensionality at which recall emerges with high probability obeys

$$d_\Psi^* \approx \frac{2 \log N}{1 - \rho^2}.$$

Intuition. The inner products $\langle \psi_i, \hat{c} \rangle$ follow a Gaussian concentration around zero with variance $1/d_\Psi$. As d_Ψ increases, the chance of spurious overlap decays exponentially, yielding a phase transition in retrieval selectivity. Hence, successful recall requires that the contextual manifold's dimensionality exceed the interference bound determined by N and mean similarity ρ. $\square$

6 Simulation Studies

To illustrate the generality of the unified operator framework and to connect it with existing empirical findings in memory research, we conducted a series of simulations. These simulations do not rely on any new behavioral data, but instead demonstrate how variations in operator placement, context regime, and retrieval order reproduce a wide range of classic memory phenomena.

6.1 Simulation 1: Operator Placement and Forms of Forgetting

The first simulation examined how apparent differences between decay, interference, and distinctiveness effects can be expressed as operator placement within the same formalism. We encoded a list of T items according to

$$W = \sum_{t=1}^{T} \gamma_t \, f_t \otimes \psi_t,$$

where f_t is the item feature vector, ψ_t is the context vector, and γ_t the encoding strength. We then introduced four distinct operators applied at different levels:

(a) **Context Compression:** $\psi'_{\text{test}} = \lambda \psi_{\text{test}}$, producing temporal compression analogous to SIMPLE.

(b) **Context Drift:** $\psi'_{\text{test}} = \rho \psi_{\text{prev}} + \eta$, corresponding to TCM-like contextual reinstatement.

(c) **Encoding Decay:** $\gamma'_t = \gamma_t e^{-\kappa \cdot \text{lag}_t}$, modeling strength-based forgetting.

(d) **Trace Decay:** $W' = (1 - \delta)W$, representing representational loss.

Despite differing in locus, all four operators generated qualitatively similar serial-position gradients, confirming that many traditional "theories of forgetting" correspond to shifts in operator placement rather than distinct mechanisms.

6.2 Simulation 2: Continuum Between Context Regimes

We next simulated a continuous transition among the three context regimes outlined in the theoretical framework: product-space (Regime 1), embedding-space (Regime 2), and joint-embedding (Regime 3). A single generator parameter $\alpha \in [0, 1]$ governed the interpolation:

$$\psi_t(\alpha) = (1 - \alpha)e_t + \alpha(\rho \psi_{t-1} + \eta f_t),$$

where e_t is a positional basis vector. Varying α smoothly transformed the model from position-based to association-based dynamics. Lag-conditional response probability (lag-CRP) analyses showed the gradual emergence of forward contiguity as α increased, demonstrating that these regimes form a continuous manifold within a single operator family.

6.3 Simulation 3: Recognition and Recall as Operator Reordering

The unified operator allows recognition and recall to be realized as different orderings of the same transformations. Recognition computes a scalar familiarity via

$$a(i \mid \text{cue}) = f_i^\top W \psi_{\text{cue}},$$

while recall adds a gating projection on context prior to feature evaluation:

$$a(i \mid \text{cue}) = \begin{cases} f_i^\top W \psi_{\text{cue}}, & \text{if } \|\psi_i - \psi_{\text{cue}}\| < \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

By increasing the sharpness parameter β of a Gaussian context kernel, recognition behavior transitioned smoothly into recall-like one-to-one retrieval. This demonstrates that recall and recognition emerge from the same operator under different context-resolution constraints.

6.4 Simulation 4: Retrieval-Driven Context Updating and Test-Position Effects

To account for the empirical pattern observed in our recognition experiments—rising target accuracy and falling foil accuracy across test positions—we introduced retrieval-driven context updating:

$$\psi_{k+1} = \rho\psi_k + \eta f_{\text{retrieved}},$$

where $f_{\text{retrieved}}$ is either the recalled or probed feature vector. This recursive updating causes the context to drift toward recently tested items, enhancing subsequent matches for true targets while reducing matches for unrelated foils. The resulting pattern matches the empirical reversal relative to standard REM predictions, suggesting that such effects emerge from retrieval-context evolution rather than encoding differences.

6.5 Simulation 5: Context Dimensionality and the Recall Threshold

Finally, we examined the dimensionality requirement for recall by varying the context dimension d_Ψ. For each simulated dataset, we computed the probability that the context-projection gate yielded exactly one matching trace. A sharp transition emerged: below a critical d_Ψ, recall collapsed into recognition-like diffuse activation. This quantifies the theoretical claim that successful recall requires a high-dimensional and sharply weighted context manifold.

6.6 Summary of Simulation Results

Across all simulations, a single operator form reproduced qualitative signatures of distinct memory models: (1) operator placement explained multiple forms of forgetting; (2) varying α unified positional and associative models; (3) recognition and recall emerged as continuous transformations; (4) retrieval-driven context drift accounted for reversed serial-position patterns; and (5) recall failures followed from low-dimensional context resolution. Together, these findings substantiate the framework's claim that many apparent theoretical divisions in memory modeling can be interpreted as reparameterizations within a single operator family.

7 Data and Methods

8 Results

8.1 *Tables and Figures*

8.2 Equations

References