A dynamic model of context-based retrieval

Madison D. Paron* James D. Paron† Michael J. Kahana‡

August 27, 2024

Abstract

We propose a comprehensive model of how experiences are encoded and retrieved from memory. At the core of the model is a dynamic retrieval process incorporating two essential mechanisms: iterative retrieval, whereby information is sequentially sampled from memory to access the full history of experiences; and competitive retrieval, whereby the most prominent features in memory inhibit the recollection of other features. Together with context-based encoding, the model quantitatively explains well-known facts about response order and inter-response times in recall experiments. We show that our retrieval process maps closely to existing decision frameworks, such as drift-diffusion models, suggesting that the memory system plays a fundamental role in a wide-ranging set of decision-making settings.

Keywords: Memory, Recall, Decisions, Context, Response times

*Department of Psychology, University of Pennsylvania. Email: mparon@sas.upenn.edu.
†Department of Finance, University of Pennsylvania. Email: jparon@wharton.upenn.edu.
‡Department of Psychology, University of Pennsylvania. Email: kahana@psych.upenn.edu.

1 Introduction

Consider an individual faced with a decision-making problem. To obtain the information necessary to reach a decision, the individual typically needs to retrieve and accumulate evidence from memory. Indeed, decisions take time in large part because evidence must be sampled sequentially from memory (Shadlen & Shohamy, 2016). This insight is missing from conventional memory models (see Kahana, 2020). These models provide detailed accounts of how memory dynamics govern the encoding of information, but appeal to separate evidence-accumulation processes (e.g., Brown & Heathcote, 2008; Ratcliff, 1978; Usher & McClelland, 2001) when describing the subsequent retrieval and response. They therefore say little about *how* experiences are sampled during retrieval and used to reach a response.

This paper argues that the same iterative memory process underlying encoding also underlies retrieval and response. We incorporate this insight into a temporal-context-based model of dynamic retrieval (TCM-DR), which provides a comprehensive account of how information is not only encoded, but also retrieved from memory to reach a decision. The model features agents (subjects) who perceive and encode information from an environment and subsequently retrieve that information in response to a cue or objective. Encoding of experiences is as in the standard temporal context model (Howard & Kahana, 2002; Sederberg, Howard, & Kahana, 2008): the subject perceives features of the environment and encodes them as neural representations in memory together with the current latent mental state, or *context*. Retrieval of features from memory is modeled through a new iterative retrieval process whereby the brain sequentially samples information from memory and maps it to recognizable objects or states from the environment.

Figure 1 conveys the basic idea of our model (the right panel) and contrasts it with existing models of memory (the left panel). In standard models, the encoding, retrieval, and response processes are separated. Some such models have no temporal retrieval dynamics at all. For instance, Howard and Kahana (2002) propose a single-step retrieval process in which one set of retrieved features forms retrieval probabilities, then the subject randomly recalls items according to these probabilities. More recent context-based models do imply temporal dynamics. In Sederberg et al. (2008), for example, encoding is the result of endogenous

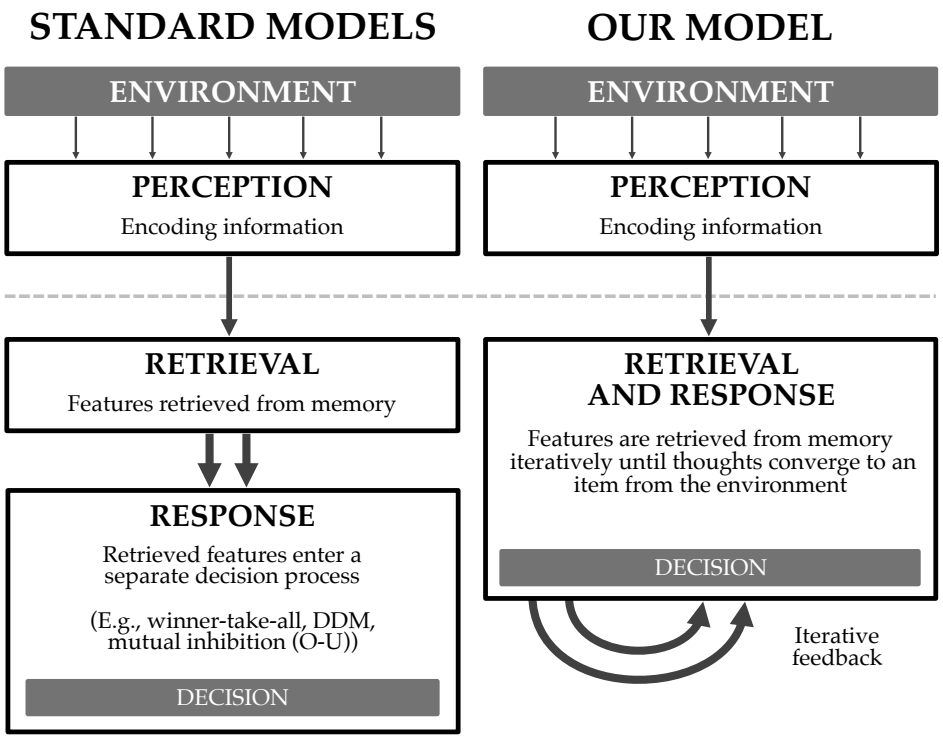

Figure 1: Comparison between the standard class of memory models and our model. Standard models separate retrieval of information in memory from the response process. Our model assumes that the same memory process underlying perception and encoding also underlies a joint retrieval and response process.

item- and context-reinstatement into memory, while retrieval and response are determined by an exogenous diffusion-based model that stochastically recalls item features. These separate response processes conceal the fact that memory is fundamentally involved not only in encoding, but also in retrieval and response. In reality, a large part of the exogenous noise driving these models is standing in for a sequential process whereby evidence is accumulated from memory.

In TCM-DR, there is a single, iterative retrieval and response process that follows the same dynamics as the encoding process. Each successful retrieval of a feature is the outcome of a sequential process in which information is sampled from memory over time (Figure 1). The subject follows a "train of thought" that is initiated by the current mental context and any relevant external cues. Features retrieve contexts, then these contexts retrieve features, and so on. Information is accumulated until a decision-relevant response is reached. There is no exogenous noise in the retrieval and response process.

We find that two psychological mechanisms are essential to explaining the dynamics

of retrieval and response: iterative retrieval and competitive retrieval. Iterative retrieval is a recursive process by which the brain uses context to retrieve features and features to retrieve contexts until these retrieved features (thoughts) become stable. It is in line with behavioral and neurological evidence that decisions take time precisely because individuals sequentially sample information from memory (Shadlen & Shohamy, 2016). We show that iterative retrieval allows the subject to access the full historical distribution of observed features. Therefore, it is helpful for forming a complete database and for generalizing beyond the present context. It is insufficient on its own, however, for two reasons: (i) it implies context- and cue-independent recall and (ii) it results in the recollection of an unrecognizable combination of past features instead of a single discernible feature.

Competitive retrieval resolves both of these issues. With competitive retrieval, the most prominent features brought to mind during retrieval are given outsized attention, suppressing less prominent features. This allows the "fuzzy" features retrieved from iterative retrieval to be "cleaned" and mapped back to recognizable features. This same mechanism is fundamental to a large class of models in computational cognitive science, notably Hopfield networks (Masson, 1995; Rizzuto & Kahana, 2001; Kahana, 2012). We show that, when iterative retrieval and competitive retrieval operate simultaneously, retrieved features can converge to a distinct past experience that depends on the initial context and cue. The subject is able to retrieve distinct features that were encoded in the distant past (due to iterative retrieval) but related to the present context and cues (due to competitive retrieval). This yields not only variation in which features are retrieved, but also variation in how long it takes to converge to a response. Critically, this variation arises even without any exogenous noise in the response process.

TCM-DR is not just a theory, but a quantitative account of the data. We validate the retrieval process by replicating empirical facts from free-recall experiments. The simulated model empirically explains classic recall-order phenomena like the recency, primacy, and temporal contiguity effects. Perhaps more interestingly, the model explains multiple dimensions of inter-response-time (IRT) data well. This includes the cross-subject distribution of IRTs and the average IRTs between list transitions by serial-position lag. A striking property of the model is that it can generate significant variation in response order and response

times, both across subjects and within subjects across trials, despite relying on only two core parameters (one for each retrieval mechanism) that are identical across subjects.[1] The only difference between subjects is their pre-experimental contexts and inter-item associations (memories). The reason is as just discussed: small differences in pre-experimental contexts can lead to large differences in what is ultimately recalled and in how long it takes for a particular feature to converge during retrieval. TCM-DR therefore explains, and provides a convincing basis for, the significant heterogeneity in retrieval we see across people and time.

Our framework stands among a large set of frameworks modeling how information is accumulated and used to reach a decision, including drift-diffusion models and Bayesian and statistical learning models. We believe that TCM-DR is not simply an alternative to these models, but a memory-based microfoundation that elucidates their deeper psychological origins. To show this, we compare TCM-DR with each class of model in turn, explaining the close mapping between them and providing a memory-based interpretation for these models' parameters and mechanisms. This mapping is especially close in diffusion-based models of retrieval, in which the basic functions of iterative retrieval and competitive retrieval are largely captured by mutual inhibition and memory drift. These insights suggest an important role for memory in a broader set of decision-making settings.

2 TCM-DR: A model with dynamic retrieval

2.1 Environment

We consider an agent (the *subject*) who exists within an environment over time. The subject partitions dimensions of the environment into elements y of a countable set $\mathcal{Y}$. At any given time, the subject's objective is either (i) to perceive and encode information from the environment or (ii) to retrieve past information in response to a cue. We describe how the agent pursues each of these objectives in turn.

[1] This is in contrast with standard drift-diffusion models, which assume exogenous noise and often assume stochastic drift rates that vary across subjects and trials (Ratcliff & Tuerlinckx, 2002).

2.2 Encoding with temporal context

Encoding is the process by which the subject stores observations of experienced features and the associations formed between them. Most of this process follows the temporal context model of Howard and Kahana (2002) and its successors (e.g., Sederberg et al., 2008).

For ease of exposition, let us index the features $y_i \in \mathcal{Y}$ by $i \in \{1, \ldots, n\}$. When the subject perceives feature y_i, it evokes a feature representation $f_i \in \mathbb{R}^n$. Each f_i is a standard basis vector, so that these elementary features are orthonormal.[2] Complex objects (e.g., a yellow square) are composites of these elementary features. At any given time t, the latent mental state of the subject's brain is represented by an m-dimensional unit vector c_t, which we call context. Each element of context represents neural activity, like the activation of a synapse or a network of neurons. Memory is represented by an $m \times n$ matrix M of associations between features and contexts. The rows of the memory matrix correspond to the dimensions of context, while columns correspond to dimensions of features; therefore, the (j, i)th element of M represents the strength of the association between context element j and feature y_i. This means that column i of M codes the contextual (neural) pattern that the brain uses to represent that feature. As we will see, distinct features become associated with each other when their contextual representations in M become more similar — that is, when they form links in the brain.

This way of modeling the encoding process is not conceptually different from the standard temporal context model (Howard & Kahana, 2002; Sederberg et al., 2008). We do, however, make one slight mathematical modification by assuming that vector lengths are taken with respect to the L^1 norm (as opposed to L^2), as in Wachter and Kahana (in press). This is not essential to the psychological intuition, but does yield three benefits. First, it makes the mathematical dynamics of retrieval more tractable. Second, it means there is only one memory matrix used for both encoding and retrieval.[3] Third, it means that a non-negative unit vector can naturally be interpreted as a vector of probabilities, a fact later shown to

[2] This assumption does not mean features can have no innate similarity; semantic similarity between attributes will show up as shared contextual representations in memory.

[3] In the classic temporal context model, feature-to-context and context-to-feature associations are stored separately in the matrices M^{FC} and M^{CF}, respectively. Because we use the L^1 norm, M^{FC} is the transpose of M^{CF}, so we need only keep track of one matrix.

link our model to Bayesian learning and inference and statistical decision theory.

Now consider a sequence of experienced features y_t over time $t \in \mathbb{N}$. In the model, the subject starts the encoding of this sequence of experiences with some pre-existing context c_0 and memory M_0.[4] The encoding process then occurs recursively. The observation of feature $y_t = y_i$ evokes $f_t = f_i$, which cues an input context from memory according to

$$c_t^{\text{in}} = M_{t-1}(\Gamma_{t-1}^{\text{col}})^{-1} f_t, \tag{1}$$

where $\Gamma_{t-1}^{\text{col}}$ is an $n \times n$ diagonal matrix with jth diagonal element equal to the sum of the elements in the jth column of M_{t-1}. The matrix $\Gamma_{t-1}^{\text{col}}$ serves only to normalize the columns of M such that c^{in} remains a unit vector. Input context c^{in} retrieves past contexts experienced contemporaneously with features similar to those of item i. This input then updates retrieved context c_t, the time-t state of the brain. Retrieved context evolves as a vector autoregression with rate of decay $\zeta \in [0, 1]$:

$$c_t = (1 - \zeta)c_{t-1} + \zeta c_t^{\text{in}}. \tag{2}$$

As long as ζ is less than one, context is autocorrelated, so that temporally contiguous features are stored with similar contexts. The association between this context and the observed features are then re-encoded into memory as

$$M_t = M_{t-1} + \phi_t c_t f_t^{\top}. \tag{3}$$

Because $f_t = f_i$ is the ith standard basis vector, this simply means c_t is added to the ith column of the memory matrix. Before being added to memory, this new association is scaled by the coefficient $\phi_t \geq 1$, which allows for the possibility that certain salient positions in the sequence of features may be encoded more strongly. In the standard temporal context model of free recall, this coefficient is the primacy gradient, taking the form

$$\phi_t = 1 + \bar{\phi}_0 e^{-\bar{\phi}_1(t-1)}. \tag{4}$$

[4] In a free-recall experiment, for instance, M_0 will be a subject-specific matrix of pre-experimental semantic associations between words.

It increases the strength of stored associations for earlier list items, declining from $1 + \bar{\phi}_0$ at the first item down toward 1 over the study phase.

After the feature y_t is encoded, the subject observes y_{t+1} and the process repeats.

2.3 Retrieval and response process

Retrieval is the process by which features y_i are brought to mind. Using current context and any accessible features (i.e., cues), the subject searches the memory matrix M to retrieve information. Re-initializing time to 0, the subject has an initial context vector c_{-1} and a memory matrix M_0. If there is a cue y_{cue}, then this cue elicits an updated context:

$$c_0 = (1 - \zeta)c_{-1} + \zeta c^{\text{in}}_{\text{cue}}, \quad \text{where} \quad c^{\text{in}}_{\text{cue}} = M_0(\Gamma^{\text{col}}_0)^{-1} f_{\text{cue}}.$$

If there is no cue, then the subject initiates recall from the current mental context: $c_0 = c_{-1}$. This retrieved context c_0 then retrieves a set of features $\tilde{f}^{\text{in}}_0$ from memory:

$$\tilde{f}^{\text{in}}_0 = M_0^{\top}(\Gamma^{\text{row}}_0)^{-1} c_0, \tag{5}$$

where Γ^{row} is analogous to Γ^{col} but with row sums of M (again, this is simply to normalize the output $\tilde{f}^{\text{in}}$). This retrieval step will yield high weight on those features with (i) high contextual similarity with c_0 and (ii) high historical frequency of observation. These features $\tilde{f}^{\text{in}}_0$ are then transformed into a vector f^{in}_0 by a competitive-retrieval step described below.

As Section 3 will explain in detail, it is not sufficient for the subject to retrieve features in one step. Rather, retrieval requires the sequential sampling of information from memory. In the model, the subject can do this by iteratively retrieving new features and contexts. We call this process *iterative retrieval*. We assume that, when the brain is in a retrieval state, intermediate feature information is not re-encoded into memory.[5] This means that the matrix of associations remains fixed at $M_t = M_0$. The set of equations underlying retrieval is essentially the same as in the encoding phase. The only difference is that, instead of retrieving a context from features of the environment f, the subject uses the retrieved

[5]This assumption is consistent with recent research showing that encoding or retrieval (but not both) can happen at a given time (Long & Kuhl, 2019).

features f^{in} at each iteration step. Thus, as in equations (1) and (2), the subject updates context according to

$$c_t^{\text{in}} = M_0(\Gamma_0^{\text{col}})^{-1} f_{t-1}^{\text{in}}, \quad c_t = (1-\zeta)c_{t-1} + \zeta c_t^{\text{in}}. \tag{6}$$

This new context then retrieves new features:

$$\tilde{f}_t^{\text{in}} = M_0^{\top}(\Gamma_0^{\text{row}})^{-1} c_t. \tag{7}$$

Retrieved features $\tilde{f}^{\text{in}}$ will be linear combinations of elementary basis features f_i; hence, they will not, in general, clearly correspond to any $y_i \in \mathcal{Y}$. This motivates a second step by which $\tilde{f}_t^{\text{in}}$ becomes f_t^{in}.

In order to retrieve a recognizable feature $f^{\text{in}} \approx f_i$, the brain will need to implement a "cleaning" procedure by which it maps the "fuzzy" features $\tilde{f}^{\text{in}}$ to standard basis vectors. This is a common process in models of cognitive psychology, like, for instance, Hopfield networks (Masson, 1995; Rizzuto & Kahana, 2001; Kahana, 2012). We call this *competitive retrieval* because it a process whereby larger elements of the vector $\tilde{f}^{\text{in}}$ suppress the recollection of smaller elements. Mathematically, there is a competitive-retrieval function $F : \mathbb{R}^n \to \mathbb{R}^n$ generating retrieved features as

$$f_t^{\text{in}} = F(\tilde{f}_t^{\text{in}}). \tag{8}$$

The function F maps the unit vector $\tilde{f}^{\text{in}}$ into a unit vector that is closer to a standard basis vector. We give F the following properties:

1. F is differentiable at each element of the input vector $\tilde{f}^{\text{in}}$.

2. F increases the larger values of $\tilde{f}^{\text{in}}$ and decreases the smaller values of $\tilde{f}^{\text{in}}$: for any i and i' such that $\tilde{f}^{\text{in}}(i) > \tilde{f}^{\text{in}}(i')$, $f^{\text{in}}(i) - \tilde{f}^{\text{in}}(i) > f^{\text{in}}(i') - \tilde{f}^{\text{in}}(i')$.

The first assumption is for analytical convenience. The second implies that the largest element of $\tilde{f}^{\text{in}}$ increases and, to keep f^{in} at unit length, the smallest element decreases. An

example of a function F satisfying these properties is the power rule:

$$f_t^{\text{in}} = F(\tilde{f}_t^{\text{in}}) = \frac{(\tilde{f}_t^{\text{in}})^\eta}{||(\tilde{f}_t^{\text{in}})^\eta||}, \tag{9}$$

where $\eta > 1$, $||\cdot||$ is the length of a vector under the L^1 norm, and exponentiation is an element-wise operation.[6] (When $\eta = 1$, we revert back to a case without competitive retrieval, in which $f^{\text{in}} = \tilde{f}^{\text{in}}$.) While our results do not depend on this specific functional form of competitive retrieval, we will continue to use it to model this step throughout the rest of the paper.

This set of equations defines how the subject retrieves f_t^{in} from f_{t-1}^{in}, using context as an intermediate step. This recursive process iterates forward on the retrieved features until a response rule is satisfied. The response rule determines (i) the point at which the iteration described above should stop; and (ii) whether, upon stopping, the retrieved feature vector is recognizable as some f_i (i.e., some $y_i \in \mathcal{Y}$). We assume two main criteria. First, retrieved features must converge to some stable vector f_∞^{in}. In particular, the process must reach some point at which $||f_t^{\text{in}} - f_{t-1}^{\text{in}}|| < \varepsilon$ for some small threshold $\varepsilon > 0$. Second, the convergent retrieved features must be sufficiently close to a feature from the environment (a standard basis vector) that it is recognizable enough to be recalled. This means that the maximum element of f_t^{in} must exceed some threshold f^{thresh} close to 1. If and once these two criteria are satisfied, the subject recalls the item corresponding to the maximum element of f_t^{in}; if the criteria are not satisfied after a given decision-time limit, nothing is recalled.

In summary, a response is completed either when convergence occurs or when some time limit is reached. To re-initiate the retrieval process and recall another feature, the subject uses the current context and an environmental cue (if available) to re-start the process described above.[7] The feature cue may be the most recently recalled item (e.g., in free recall) or might be an attribute of the initial cue.

[6] That is, the operation takes the power of each individual element: for some arbitrary vector x, the ith element of x^η equals $x(i)^\eta$.

[7] Additional retrieval and response criteria may be added depending on the situation. For example, in a free-recall task, one usually assumes that the subject will suppress already-recalled features in M, preventing repetition. We implement this in the application in Section 4.

3 Dynamics of the retrieval process

Having laid out the assumptions of the model, we now analyze the underlying mechanisms and their consequences for retrieval over time. We do this by mathematically characterizing and discussing the dynamics of the retrieval induced by iterative retrieval and competitive retrieval. These properties are transparently illustrated by way of a simple two-item example.

3.1 Evolution of retrieved features

By combining the equations constituting the retrieval process, one can express the temporal dynamics of retrieved features f_t^{in} autoregressively. This eliminates the need to keep track of intermediate latent contexts. In particular, substituting context evolution (6) and recalled features (7) into the expression (8), we have the relation

$$f_t^{\text{in}} = F\left((1-\zeta)\tilde{f}_{t-1}^{\text{in}} + \zeta\Phi f_{t-1}^{\text{in}}\right), \tag{10}$$

where the matrix

$$\Phi = M_0^\top (\Gamma_0^{\text{row}})^{-1} M_0 (\Gamma_0^{\text{col}})^{-1}. \tag{11}$$

This is a feature-to-feature transition matrix that summarizes the mapping from one set of retrieved features, to an intermediate retrieved context, back to a new set of retrieved features. In other words, Φ summarizes the transition rates between features of the environment based on their shared historical associations.

What does the dynamical system (10) tell us about how retrieved features evolve over time? We can first answer this by decomposing the changes implied by (10) into its constituent mechanisms. Letting $\Delta\tilde{f}_t^{\text{in}}$ denote the temporal difference $\tilde{f}_t^{\text{in}} - \tilde{f}_{t-1}^{\text{in}}$, we have

$$\Delta\tilde{f}_t^{\text{in}} = \underbrace{\zeta(\Phi - I)f_{t-1}^{\text{in}}}_{\text{iterative retrieval}} + \underbrace{\zeta(f_{t-1}^{\text{in}} - \tilde{f}_{t-1}^{\text{in}})}_{\text{competitive retrieval}}. \tag{12}$$

The first term in this expression is the effect of iterative retrieval. Multiplication with the transition matrix Φ represents a sampling of experiences, through contextual associations, using the features currently in mind. On its own, it appears as a standard linear dynam-

ical system (i.e., a vector autoregression). The second term of (12) represents competitive retrieval. By assumption, the function $F(\tilde{f}^{\text{in}})$ increases large elements of $\tilde{f}^{\text{in}}$ and decreases small elements. Thus, the difference $f^{\text{in}} - \tilde{f}^{\text{in}}$ is strictly positive for the largest element and strictly negative for the smallest, summarizing the change induced by F. Together, these effects pull retrieved features in two directions: toward those that are most easily sampled from memory and toward those that are currently most prominent.

To get a more direct interpretation of the evolution of f^{in}, we can calculate the change in actual retrieved features Δf^{in}, as opposed to the intermediate step $\Delta \tilde{f}^{\text{in}}$. Appendix A.1 proves that Δf^{in} can be expressed in the form

$$\Delta f_t^{\text{in}} = \underbrace{D^F(\tilde{f}_{t-1}^{\text{in}})\zeta(\Phi - I)f_{t-1}^{\text{in}}}_{\text{iterative retrieval}} + \underbrace{D^F(\tilde{f}_{t-1}^{\text{in}})\zeta(f_{t-1}^{\text{in}} - \tilde{f}_{t-1}^{\text{in}})}_{\text{competitive retrieval}} + o_{t-1}, \tag{13}$$

where the term o_{t-1} represents small terms that disappear as the time interval between iterations approaches zero (i.e., the continuous-time limit).[8] The matrix D^F is an $n \times n$ Jacobian matrix with elements

$$D_t^F(i, i') = \frac{\partial F(\tilde{f}_t^{\text{in}}(i))}{\partial \tilde{f}_t^{\text{in}}(i')}.$$

This Jacobian D^F summarizes the non-linearity induced by the transformation F. Indeed, noting that

$$\Delta f_t^{\text{in}} = D^F(\tilde{f}_{t-1}^{\text{in}})\Delta \tilde{f}_t^{\text{in}} + o_{t-1},$$

we can see that updating occurs effectively in two separate parts. Starting with a prior set of retrieved features $\tilde{f}_{t-1}^{\text{in}}$, competitive and iterative retrieval first yield an intermediate set of retrieved features $\tilde{f}_t^{\text{in}}$, as described above. Then, these changes are mapped toward a unit vector to complete the competitive retrieval step. The matrix D^F summarizes the directions in which each component travels to accomplish this.[9]

[8] Equivalently, the term o_{t-1} represents second- and higher-order terms in a Taylor expansion of f_t^{in}.

[9] Appendix A.2 derives the explicit form of D^F under our power rule with parameter η.

3.2 Understanding the mechanisms

Iterative retrieval and competitive retrieval each play essential roles in shaping the dynamics of retrieved features. To show this, we examine each in isolation. First, we shut down competitive retrieval and study the effects of iterative retrieval; second, we shut down iterative retrieval and study competitive retrieval.

3.2.1 Understanding iterative retrieval

Suppose there is no competitive retrieval: $f^{\text{in}} = \tilde{f}^{\text{in}}$. Retrieved features evolve according to the simple vector autoregression (VAR):

$$\Delta f_t^{\text{in}} = \zeta(\Phi - I) f_{t-1}^{\text{in}}. \tag{14}$$

This is a linear dynamical system; to understand whether (and to what) it converges, we need to study the properties of the transition matrix Φ. Note first that, iterating (14) backward to time 1, we can rewrite the system as an explicit function of time:

$$f_t^{\text{in}} = A^t f_0^{\text{in}}, \quad \text{where} \quad A = (1-\zeta)I + \zeta\Phi \quad \text{and} \quad f_0^{\text{in}} = M_0^\top (\Gamma_0^{\text{row}})^{-1} c_0. \tag{15}$$

Establishing unique convergence of this system amounts to studying the eigenvalues of A. Note first that, for an n-dimensional vector of ones ι_n, we have

$$\iota_n^\top \Phi = \iota_n.$$

That is, the columns of Φ all sum to one, and ι_n is a left eigenvector of Φ with corresponding eigenvalue equal to one. From (15), the same must be true of the matrix A. The elements of Φ and A are strictly positive, so these are regular, column-stochastic matrices. It follows from standard linear algebra results that Φ and A each have a well-defined right-eigenvector corresponding to the eigenvalue of one. All other eigenvalues are strictly less than one, so this system does indeed converge to a fixed point.

More specifically, Appendix B proves that the dynamical system (15) always converges

to a positive unit vector f_∞^{in}, which is the unique solution to the linear system

$$f_\infty^{\text{in}} = \Phi f_\infty^{\text{in}}. \tag{16}$$

This result not only guarantees that the solution is unique, but also reveals that it is a function of the memory matrix M only. The initial context c_0 with which the subject initiates recall is irrelevant to the convergence of the retrieval process. We can think of Φ as revealing ergodic probabilities of encoded features: the long-run historical frequencies with which certain items have been observed over time. Thus, iterative retrieval on its own represents a process by which the brain sequentially samples information from memory until it obtains a sufficiently representative sample from history.

This context-independent property of iterative retrieval is ideal for recalling as much historical information encoded in memory as possible; however, it has two obvious drawbacks. First, it fails to map retrieved features f^{in} to items from the environment. It leads instead to combinations of item features distributed according to their historical frequencies. Second, it eliminates the ability of the brain to use cues and temporal context as important conditioning information. In most decision-making settings, one should presumably rely on context. Take as an example a free-recall task. Context in the recall phase allows the subject to recall items in the most recent list, as opposed to items in other lists or outside of the experiment. Context also allows the subject to sequentially cue contiguous words in the list, recognizing that they are in some way associated. Competitive retrieval resolves both of these issues.

3.2.2 Understanding competitive retrieval

Let us now shut down iterative retrieval and study the role of competitive retrieval. Again, we can express retrieved features as a function of time only:

$$f_t^{\text{in}} \propto (f_0^{\text{in}})^{\eta t}, \quad \text{where} \quad f_0^{\text{in}} = M_0^\top (\Gamma_0^{\text{row}})^{-1} c_0, \tag{17}$$

where, as before, the exponent is taken element-by-element. Provided that the initial set of features f_0^{in} has a unique maximum value, this expression will converge to a unit vector f_∞^{in}

such that[10]

$$f_{\infty}^{\text{in}}(i) = \begin{cases} 1 & \text{if } f_0^{\text{in}}(i) > f_0^{\text{in}}(i'),\ \forall i' \neq i, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

The most prominent feature brought to mind by initial cues and context "wins" the competition for retrieval.

In contrast to iterative retrieval, which yields a set of retrieved features that are entirely independent of the initial cues and context, competitive retrieval yields a set of features that are solely a function of initial cues and context. This initial context recalls a set of features from memory, and the brain then maps those features to items from the environment at a rate governed by η. As already mentioned, this process is analogous to the decoding of noisy information in a Hopfield network. Using available information in context c_0, the brain decodes from memory a noisy representation f_0^{in} of an item y_i from the environment $\mathcal{Y}$. It then processes this noisy representation via a "cleaning" process to render it recognizable; this process takes time, so that recognition of the initial representation is not immediate.

3.2.3 Understanding the combined effects

Iterative retrieval and competitive retrieval operate together, resulting in a recall process that depends on both long-run historical associations and the initial cues and context at the time of recall. Because of the non-linearities in this joint process, the system characterizing recall dynamics appears much more complicated. Still, we can show that it has stable convergence properties and that the ultimate point of convergence depends on the point of recall initiation.

Combining both mechanisms means that the dynamical system (10) converges to a positive unit vector f_{∞}^{in}, which is one of the solutions to the nonlinear system

$$f_{\infty}^{\text{in}} = F(\Phi f_{\infty}^{\text{in}}). \tag{19}$$

This system may have multiple solutions due to the non-linearity of the function F. Under

[10] In the unusual, knife-edge case when f_0^{in} has $N \geq 2$ identical maxima, the corresponding elements of f_{∞}^{in} become $1/N \leq 1/2$. Under our decision rule f^{thresh}, this would be too small to recall any value, so nothing would be recalled.

our power rule, the expression (19) can be rewritten as the system

$$f_\infty^{\text{in}}(i)\left(\sum_{i=1}^{n}\left(\sum_{i'=1}^{n}\Phi(i,i')f_\infty^{\text{in}}(i')\right)^{\eta}\right)=\left(\sum_{i'=1}^{n}\Phi(i,i')f_\infty^{\text{in}}(i')\right)^{\eta}$$

for every $i \in \{1,\ldots,n\}$. Expanding these sums and products results in a polynomial system with n equations in n unknowns. As this polynomial is of degree greater than one (because $\eta > 1$), the solution need not be unique as in the linear ($\eta = 1$) case.

The particular solution to which this system converges over time, then, depends on the initial position f_0^{in}, which itself is a function of the initial context c_0. Different initial contexts can yield dramatically different convergent retrieved features. Another way in which to view this is using the dynamics (13), which shows that the evolution of retrieved features is the net effect of an iterative retrieval term and a competitive retrieval term. Iterative retrieval moves current features in the direction implied by Φ, as described in Section 3.2.1. Competitive retrieval moves features toward a unit vector. The sum of these two effects determines not only the final vector to which the process will converge, but also the response time. If these effects move retrieved features in the same direction, convergence will be faster; if, instead, they work in opposite directions, convergence can be very slow.

3.3 Illustration: A two-item setting

To illustrate the convergence properties of our retrieval and response process more concretely, we examine it in a setting with only two items ($n = 2$). The example has the added benefit of helping to interpret our model in light of the common scenario in which an individual must choose between two given alternatives. The two-alternative forced choice (TAFC) task is one classic experimental example of this (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006).

For simplicity, suppose that the memory matrix M_0 takes the symmetric form

$$M_0 \propto \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}$$

for some $p \in (0,1)$. This means that the two item features have been experienced with

equal historical frequency (have equal column sums) and have some shared associations ($p \notin \{0, 1\}$). Assume $p \neq 1/2$, so that these items do not have trivially identical representations in memory. This memory matrix implies a symmetric transition matrix

$$\Phi = \begin{bmatrix} 1-\phi & \phi \\ \phi & 1-\phi \end{bmatrix},$$

where $\phi = 2p(1-p) \in (0, 1/2)$. Notice that the ergodic probability vector for Φ — the vector f_∞^{IR} satisfying $f_\infty^{\text{IR}} = \Phi f_\infty^{\text{IR}}$ — is equal to

$$f_\infty^{\text{IR}} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}.$$

We use the "IR" superscript to emphasize that this is the solution to the model with only iterative retrieval ($\eta = 1$). Absent competitive retrieval, all initial cues and contexts result in this set of retrieved features.

Because $f_t^{\text{in}}(1) = 1 - f_t^{\text{in}}(2)$, we can describe the whole system with just one element; consider the first element $f_t^{\text{in}}(1)$. From the fixed-point formula (19), we have that

$$f_\infty^{\text{in}}(1) = \frac{[\phi + (1-2\phi) f_\infty^{\text{in}}(1)]^\eta}{[\phi + (1-2\phi) f_\infty^{\text{in}}(1)]^\eta + [1 - (\phi + (1-2\phi) f_\infty^{\text{in}}(1))]^\eta}. \tag{20}$$

In the special case of $\eta = 2$, for example, this fixed-point equation reduces to a cubic function of $f_\infty^{\text{in}}(1)$, to which there are three unique and distinct solutions. One solution is $f_\infty^{\text{in}}(1) = 1/2$, the knife-edge case in which the memory-iteration and competitive-retrieval solutions are the same. The other two solutions are symmetrically above and below $1/2$: one is in the interval $(0, 1/2)$ and the other is in $(1/2, 1)$. Thus, these two final retrieved features vectors lie somewhere between the ergodic vector with $f_\infty^{\text{in}}(1) = 1/2$ decoded from memory and the extreme binary outcomes $f_\infty^{\text{in}}(1) = 0$ or $f_\infty^{\text{in}}(1) = 1$ from competitive retrieval.

How do retrieved features evolve between these fixed points? The difference equation

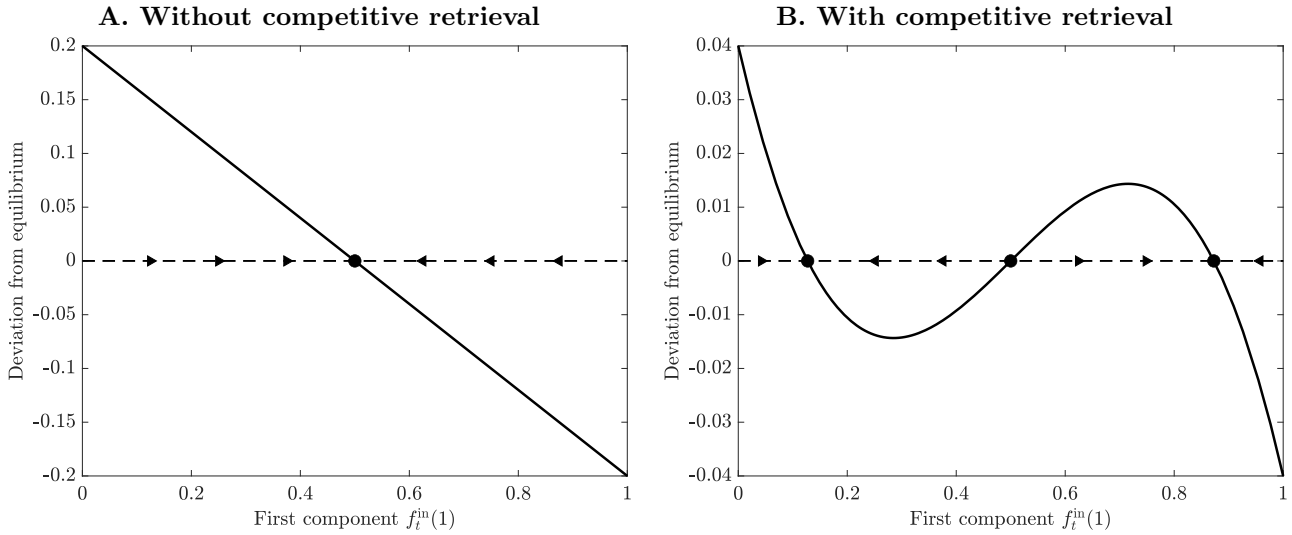

Figure 2: Phase diagrams for the two-item example with and without competitive retrieval. The solid lines are the deviation of $f_t^{\text{in}}(1)$ from a convergent fixed point, as defined in (21). Each intersection of these lines with zero represents a fixed point $f_\infty^{\text{in}}(1)$. Arrows along the zero line represent the direction in which $f_t^{\text{in}}(1)$ moves over time given that initial position. Panel A assumes $\eta = 1$ and Panel B assumes $\eta = 2$. Both plots assume $\phi = 0.2$.

(12) implies that, from $t-1$ to t, the change in the first element of $\tilde{f}^{\text{in}}$ equals

$$\Delta \tilde{f}_t^{\text{in}}(1) = \underbrace{\zeta\phi(1 - 2f_{t-1}^{\text{in}}(1))}_{\text{iterative retrieval}} + \underbrace{\zeta(f_{t-1}^{\text{in}}(1) - \tilde{f}_{t-1}^{\text{in}}(1))}_{\text{competitive retrieval}}.$$

The iterative retrieval term is strictly positive for all $f_{t-1}^{\text{in}}(1) < 1/2$ and strictly negative for $f_{t-1}^{\text{in}}(1) > 1/2$. The opposite is true of the competitive-retrieval term: it is positive for $f_{t-1}^{\text{in}}(1) \in (1/2, 1)$ and negative for $f_{t-1}^{\text{in}}(1) \in (0, 1/2)$. At $f_{t-1}^{\text{in}} = 1/2$, these effects are both zero, because $f_t^{\text{in}} = f_\infty^{\text{IR}}$ and both elements equal the maximum of f_t^{in}. Iterative retrieval dominates for $f^{\text{in}}(1)$ near 0 or 1 and competitive retrieval dominates for $f^{\text{in}}(1)$ near 1/2.

Figure 2 summarizes the dynamics of the retrieval process in this two-item example via a phase diagram. Panel A corresponds to the $\eta = 1$ case; Panel B corresponds to the $\eta > 1$ case. In each case, the solid line represents the deviation of current retrieved features $f_t^{\text{in}}(1)$ from the fixed-point condition (20). In particular, the deviation equals the difference

$$\text{Deviation}(f_t^{\text{in}}(1)) = \frac{[\phi + (1-2\phi)f_t^{\text{in}}(1)]^\eta}{[\phi + (1-2\phi)f_t^{\text{in}}(1)]^\eta + [1 - (\phi + (1-2\phi)f_t^{\text{in}}(1))]^\eta} - f_t^{\text{in}}(1). \quad (21)$$

Of course, (20) implies that this deviation is equal to zero at every f_{∞}^{in}. The arrows in the figures illustrate that, between these fixed points, $f_t^{\text{in}}(1)$ increases when the deviation is positive and decreases when it is negative. In the $\eta = 1$ case, regardless of the current (and therefore initial) condition, retrieved features converge over time to f_{∞}^{IR}, consistent with historical frequencies embedded in Φ. In the $\eta > 1$ case, there are three possible points of convergence. The point f_{∞}^{IR} is an unstable fixed point: retrieved features converge here if and only if the initial context is such that $f_0^{\text{in}}(1) = 1/2$. The other two points are stable: for all c_0 such that $f_0^{\text{in}}(1) < 1/2$, retrieved features converge to the low fixed point and the subject chooses item 2; and for all $f_0^{\text{in}}(1) > 1/2$, features converge to the high fixed point and the subject chooses item 1.

This two-item example is evidently stylized and hence abstracts from some of the complexities of settings with many items. As the dimensionality of the feature space (and perhaps η) becomes larger, the number of possible fixed points and the dynamics between them will change, as we will see in our quantitative evaluation of the model. This will yield a larger variance in both the possible convergence points and the time it takes to reach them. Still, the core logic remains the same. The interaction of iterative retrieval and competitive retrieval represents a trade-off among three forces: the decoding of historical information from memory, the mapping of internal representations to features from the environment, and the utilization of current context to make more relevant decisions.

4 Application to free-recall data

To validate our framework, we apply TCM-DR to the experimental setting of a free-recall experiment. Our objective with this application is to study whether the model is capable of quantitatively explaining patterns in recall order and inter-response times. To do this, we calibrate the model parameters and simulate multiple subjects studying and recalling a list of items. We find that the model simulation reproduces multiple dimensions of responses and inter-response times (IRTs) in data from a standard free-recall experiment. In particular, it is able to match the substantial within- and between-subject variation in recall behavior observed in the data, despite being fully deterministic and relying on only two core parameters

(ζ and η) that are assumed to be identical across time and subjects.

Importantly, our goal with this application is to explain as much as possible without adding additional assumptions to the model. Thus, this version of the model does not include the full set of machinery incorporated in state-of-the art temporal context models like TCM-A (Sederberg et al., 2008) and CMR (Polyn, Norman, & Kahana, 2009). For example, we do not extend our model to include the forward-asymmetry parameter (γ^{FC}) or the weighting of temporal versus source context (L^{FC}) that help these models to explain additional facts. Adding these ingredients to TCM-DR would be straightforward, but would do little to help with the validation exercise in this section.

4.1 Data and experimental methods

Data are taken from Experiment 4 of the Penn Electrophysiology of Encoding and Retrieval Study (PEERS4), a high-quality dataset comprising 682,032 spoken responses gathered from 98 subjects who each participated in 23 experimental sessions. This study involved delayed-recall of long lists (24 items), making trials with perfect recall rare. Accuracy data from this experiment are reported by Kahana, Aggarwal, and Phan (2018) and Aka, Phan, and Kahana (2021); response-time data are analyzed in detail in Greene, Goldman, and Kahana (2024).

4.2 Model calibration and simulation

Table 1 summarizes the calibration of our model and the relevant assumptions behind the simulation. As shown in Panel A, the parameters governing the encoding and retrieval processes are the rate of context updating ζ, the power rule for competitive retrieval η, and the two parameters governing primacy ϕ from (4). Primacy parameters are chosen to match recall probabilities by serial position. As we showed above, higher ζ and η will govern average inter-response times. Lower ζ will also increase the strength of temporal contiguity, as low ζ means context is more correlated during the study and retrieval phases. Higher η increases the within- and cross-agent variance of recalls and IRTs. These moments determine which parameter combination can best explain our set of empirical facts.

	Parameter	Value
Panel A: Encoding and retrieval parameters		
Rate of context updating	ζ	0.48
Competitive retrieval power rule	η	50
Primacy strength	$\bar{\phi}_0$	0.01
Primacy rate of decay	$\bar{\phi}_1$	0.5
Panel B: Decision-rule parameters		
Convergence threshold	ε	0.001
Decision threshold (before convergence)	$f_{\text{early}}^{\text{thresh}}$	0.999
Decision threshold (at convergence)	$f_{\text{conv}}^{\text{thresh}}$	0.8
Panel C: Dimensions of the experiment		
Number of list items	n	25
Number of context elements	m	25
Panel D: Cross-subject initial conditions		
Pre-experimental memory matrix	M_0^{pre}	Identity + Unif. dist.
Pre-experimental context	c_0^{pre}	Unif. dist.

Table 1: Calibration of the simulated model. See the main text for details.

Panel B of Table 1 summarizes the response-rule parameters. Recall that the retrieval and response process converges when $||\Delta f_t^{\text{in}}||$ is less than some very small number ε; we choose $\varepsilon = 0.001$. Conditional on convergence, the maximum element of f_t^{in} is chosen if and only if that element exceeds some threshold $f_{\text{conv}}^{\text{thresh}}$. We choose a threshold of 0.8. Finally, in the rare case that the process is not yet convergent but retrieved features are nearly identical to a unit vector, the subject recalls that item "early." We consider this to have happened if the maximum element of such a vector exceeds $f_{\text{early}}^{\text{thresh}} = 0.999$.[11]

Under this calibration, we can simulate the experiment for multiple subjects. Consider first Panel C of the table. We assume an experiment with $n = 25$ distinct items and a latent context with $m = 25$ different elements.[12] In our simulation, all subjects study a list of $n = 25$ words and, as Section 2 describes, subsequently attempt to freely recall as many of these items as possible, in any order. After a successful retrieval, the column of M corresponding to the just-recalled item is suppressed near zero to prevent an unreasonable

[11] We add this only because it is reasonable; it does not impact the results.

[12] The number of context elements is not particularly important, so long as this number is larger than the number of features $m \geq n$. If $m < n$, then the subject can only form a limited number of distinct inter-item associations.

amount of repetition. Strictly speaking, this suppression step is not necessary to explain recall behavior, but does reduce average inter-response times between retrievals.

Subjects differ in their initial (pre-experimental) contexts and memories. This is the only dimension along which subjects are heterogeneous; therefore, all dispersion in recall order and IRTs can be traced back to these differences. As stated in Panel D, we assume that the subject-specific initial conditions are simply drawn at random from a set of independent uniform distributions. We choose a uniform distribution to emphasize that we do not need to impose a specific kind of structural cross-subject heterogeneity to generate realistic recall behavior. The randomness of the initial contexts reflects the fact that each subject comes into the laboratory with a unique set of recent experiences.[13] The memory matrix is an equally weighted average of a common component (an identity matrix), reflecting distinctiveness of the items; and an idiosyncratic component (a uniform random matrix), reflecting differences in prior associations between words that happen to be presented on the list. Note that, because we do not specify the actual items in the list, we do not take a stand on the innate semantic similarity between list items. One could easily add that possibility to the model by assuming that, in the common component of subjects' initial memory matrices, certain items tend to share certain contexts.[14]

4.3 Results: Free-recall responses

4.3.1 Recall order

We first show that the simulated model explains classic phenomena characterizing the order in which subjects recall items. This includes the recency and primacy effects present in probabilities of first recall and the temporal contiguity effect present in lag conditional response probabilities (CRPs). The model also succeeds in explaining other facts (e.g., serial position curves). We focus on these two because they illustrate two essential properties of the model: the context-dependence of retrieval and the importance of temporal contiguity for associative learning. We assume immediate free recall throughout, but it is straightforward

[13] For instance, someone who came to the lab from the zoo will likely start with a context closer to that of the word "lion" than someone who just came from the beach.

[14] Again, item features themselves are represented by standard basis vectors, so no two items are explicitly similar in feature space; rather, semantic similarity is captured by shared contexts between items in M.

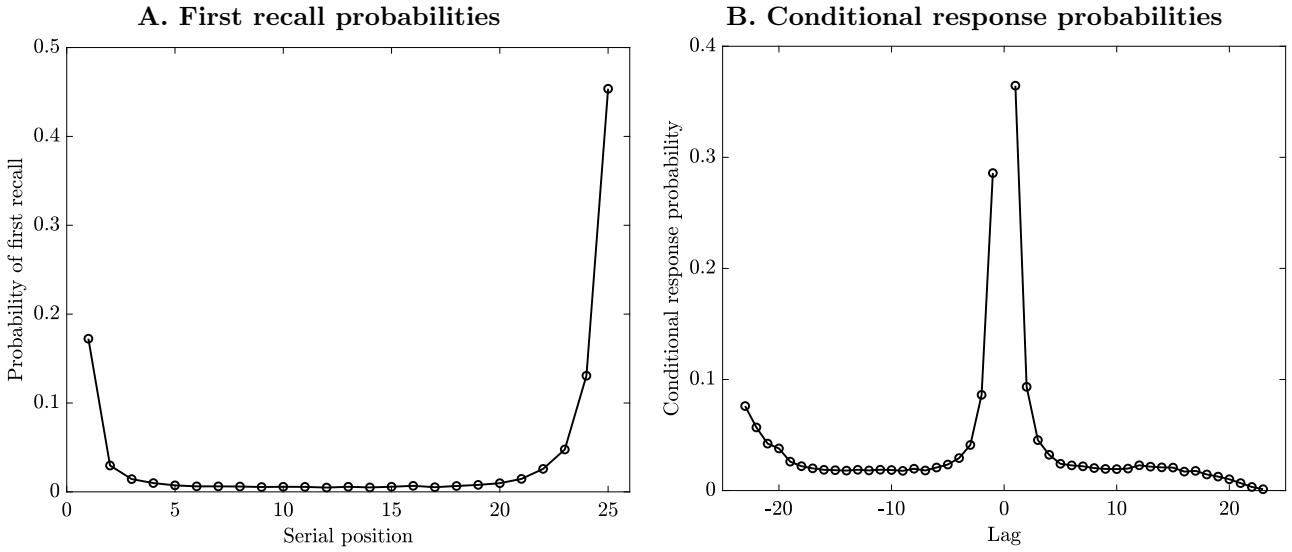

Figure 3: A. Probability of first recall in the model. Serial positions correspond to the order in which an item was presented during study. **B.** Lag conditional response probabilities (CRPs) in the model. Each lag CRP is the probability of making a transition from item i to item $i + \text{Lag}$ during recall, conditional on that transition being available. Both panels **A.** and **B.** assume immediate free recall.

to introduce a distractor into the model, which will serve to suppress recency effects.

Panel A of Figure 3 plots the model-implied probability of first recall. Because this is immediate free recall, the context used to initiate retrieval is very similar to the subject's context at the time the last list item was encoded, so the last item is very likely to be recalled first. This is the recency effect. Importantly, the most recent item is not guaranteed to be the first-recalled item: some subjects first recall items in other serial positions, and a noticeable number start at the beginning of the list (a primacy effect). In our model, small between-subject differences in pre-experimental contexts and associations lead to noticeable differences in realized responses.

Panel B of Figure 3 plots the lag CRPs. The CRP at lag j represents the probability that a subject recalls item $i+j$ after having just recalled item i, conditional on that transition being possible. The peaks at lags ± 1 characterize the temporal contiguity effect: subjects are more likely to make transitions to words that were encoded nearby in time. This is a consequence of contextual autocorrelation through the parameter ζ. There is a slight forward asymmetry, both in our model-implied CRP and in the data: subjects tend to advance forward in the list slightly more often than they move backward. The CRPs also increase for extreme negative lags, since the primacy effect increases the likelihood that subjects will transition to words

at the beginning of the list.[15] Overall, the recall probabilities and CRPs generated by the model accord well with corresponding plots in the data (e.g., Kahana et al., 2024).

4.3.2 Response times

TCM-DR is certainly not the first context-based model to explain recency, primacy, or temporal contiguity effects. It is, however, the first to do so without appealing to an exogenous one-step or diffusion-based retrieval process to generate variation in the simulated data. A more revealing test of our model is whether it can also explain variation in inter-response times (IRTs), both between subjects and within subjects across trials.[16] We find that the model explains multiple dimensions of the response-time data well, including the distribution of IRTs and the average IRTs between retrievals of varying lags. Note that our model characterizes time in terms of number of iterations, while the data are in seconds. To match average response times, we assume that 1 second in the data corresponds to 10 iterations.

Panel A of Figure 4 compares the IRT distributions in the data and model. The data suggest a highly variable, right-skewed distribution.[17] In the model, small differences in initial context lead to large differences in IRTs. Sometimes, initial context is such that the iterative retrieval and competitive retrieval effects work in the same direction, or one dominates, resulting in a fast response; other times, they work in offsetting directions so that convergence takes a long time.

Panel B of Figure 4 compares the average lag IRTs in the data and model. These curves represent the response-time analogues to the CRP curve above: each point at lag j is the average time taken between a transition from item i to item $i + j$. IRTs tend to be much shorter for close transitions, since these items tend to be encoded with similar contexts and therefore are easy to retrieve in succession. IRTs also tend to be slightly shorter at long lags (from one end of the list to the other) because of the primacy and recency effects.

[15] The model fails to produce the fact that CRPs tend to be slightly elevated for extreme positive lags. The reason for this is that we do not include the forward-asymmetry parameter in the model that is included in more recent versions of the temporal context model (e.g., Sederberg et al., 2008; Polyn et al., 2009).

[16] Drift-diffusion models typically achieve this by assuming exogenous stochastic drift rates across agents and trials (Ratcliff & Tuerlinckx, 2002); our model is able to achieve this solely by introducing heterogeneity in pre-experimental context.

[17] It is clearly non-normal, as shown by the normal fit and quantile plot in the data figure.

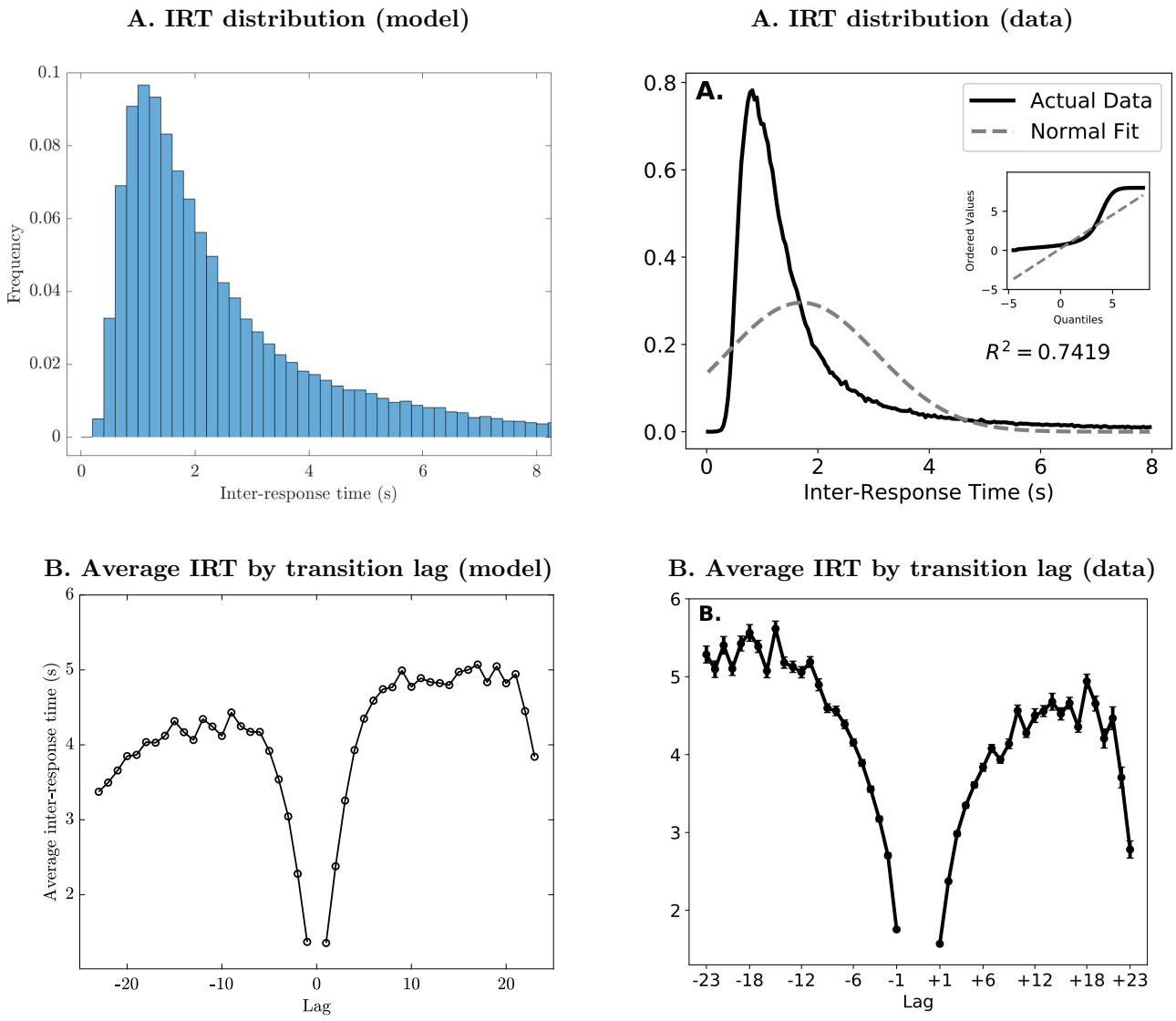

Figure 4: Inter-response times (IRTs) in the model and data. Panel A plots the distribution of IRTs in seconds (1 iteration = 0.1 seconds in the model). In the data panel (right), the normal-distribution fit and quantile plot display the non-normality of the distribution. Note that the scales of the y-axes are different for the model and data because the model panel is a histogram frequency and the data panel is a smoothed density. Panel B plots the average inter-response time for transitions between item i and item $i + \text{Lag}$.

Overall, even without any parameter heterogeneity or exogenous noise, TCM-DR succeeds at capturing response-time behavior in recall.[18]

5 Comparison with existing decision models

An important contribution of TCM-DR is that it can be used to better understand existing models of recall and decisions through the lens of the principles of memory. This section discusses how our model relates to two well-known categories of models: diffusion-based psychological models of retrieval and response and models of Bayesian and reinforcement learning.

5.1 Diffusion-based models of retrieval and choice

The iterative retrieval process at the center of TCM-DR sheds light on existing models of retrieval and choice that appeal to separate models with reduced-form drift rates and exogenous noise. As a state-of-the-art example, let us consider the leaky-accumulator model of Sederberg et al. (2008) and Polyn et al. (2009), which is based on the decision process in Usher and McClelland (2001). In this model, the subject first retrieves a set of features f_0^{in} from memory, as in (5). Instead of iteratively updating these features, though, the subject now initiates and keeps track of a new state variable x_t, an n-dimensional vector with ith element corresponding to the ith item feature. Let Δx_t denote the change $x_t - x_{t-1}$. Beginning with initial condition $x_0 = \vec{0}$, x evolves autoregressively according to the process

$$\Delta x_t = \underbrace{-\tau(\kappa + \lambda(\mathbf{L} - I))x_{t-1}}_{\text{mutual inhibition}} + \underbrace{\tau C_V f_0^{\text{in}}}_{\text{memory}} + \underbrace{\tau\sigma\epsilon_t}_{\text{noise}}, \tag{22}$$

$$x_t(i) \to \max\{x_t(i), 0\}, \quad \forall i. \tag{23}$$

This retrieval process ends in a response once one of the accumulators $x_t(i)$ reaches a threshold, at which point that item is recalled.

[18] Like the CRPs in Figure 3, the absence of the forward-asymmetry parameter of TCM-A and CMR means that we get slightly longer IRTs for positive lags than for negative lags in our model.

The first term on the right-hand side of (22) represents the mutual inhibition of item features: $\kappa > 0$ modulates the extent to which individual features inhibit themselves, while $\lambda > 0$ and $\mathbf{L}$ (a matrix filled with ones) govern how different features inhibit each other. The second term is the drift toward the initial retrieved-features vector f_0^{in}. It is scaled by the coefficient of variation C_V of f_0^{in}, a scalar which is larger when f_0^{in} is closer to a standard basis vector. This speeds up convergence when the initial feature representation is already close to an item from the environment. The third and final term is an exogenous diffusion with vector of shocks $\epsilon_t \overset{\text{iid}}{\sim} N(0, I)$ scaled by standard deviation σ. The scalar $\tau > 0$ governs the speed of convergence of the entire process.

This conventional diffusion-based model incorporates a striking number of mechanisms that arise endogenously from our memory model. We discuss these similarities between (13) and each element of the process (22) in turn. The first mechanism is mutual inhibition of item features with themselves and each other. In our context-based model, this occurs through both iterative retrieval and competitive retrieval. As we have shown above, iterative retrieval causes retrieved features to drift toward the historical distribution of features implied by memory through Φ. If a given feature element $f_t^{\text{in}}(i)$ is "too large" relative to this distribution, it will be suppressed relative to other elements. Unlike in the diffusion-based model, some features will have to grow, but this is only because we need to keep the vector on the unit circle.[19] Competitive retrieval also has an obvious mutual-inhibition-like feature in the property that large features inhibit small features.

The second mechanism in (22) is the drift term proportional to f_0^{in}. In a sense, this term constitutes a simplified view of our model. The initial retrieved-features vector f_0^{in} is the result of the first iteration of contextual retrieval; it reflects non-iterative decoding from memory. Like the competitive retrieval rule, it results in a recall process that puts weight on the initial context c_0 in the ultimate recall decision. The coefficient of variation C_V operates in much the same way as competitive retrieval via η, which more strongly directs retrieved features toward a standard basis vector if f_0^{in} is itself closer to a standard basis vector.

[19] More formally, one could conceive of an isomorphism $T : \mathbb{R}^n \to \mathbb{R}^n$ that maps unit vectors f_t^{in} to some non-unit vectors x_t. It may be that all elements of x decrease from one period to the next, even though some elements of f^{in} must increase. Because T is an isomorphism, Δf^{in} and Δx are equivalent ways of describing the retrieval dynamics.

The third component of (22) is the diffusion term $\sigma\epsilon_t$, which induces a dispersion in both choices and decision times across trials and across subjects. To an extent, our model provides a microfoundation that does not need exogenous shocks to explain this dispersion. For a given subject, cross-trial choice and decision-time variability arises naturally from changing temporal context. Between subjects, choices and decision times will vary due to differences in pre-experimental context and memory. As we have shown, the parameter η increases the variability of responses and response times in the same way as σ. The two-item example showed clearly how $\eta > 1$ increases the number of possible vectors to which f^{in} can converge, explaining the variance in retrieved features. The non-linearity introduced by competitive retrieval also leads to varying convergence speeds that depend on the initial condition.

Finally, there is an obvious correspondence between the parameter τ in the diffusion-based model and the rate of context updating ζ. When individuals update their contexts quickly, they have faster response times.[20] Similarly, the competitive-retrieval parameter η scales the updating process (the Jacobian D^F scales with η), increasing the speed of updating.[21]

5.2 Models of Bayesian and reinforcement learning

TCM-DR is a psychology-based model of learning and inference. The standard precedents for this are Bayesian learning, reinforcement learning, and statistical decision theory. We show that the encoding and retrieval processes underlying TCM-DR have a very natural relation to these frameworks. This comparison is useful for understanding the extent to which these mechanisms might be supported by principles of formal statistical learning.

We can see the Bayesian view by interpreting memory and context as arrays of probabilities.[22] Let p_t denote the subject's time-t subjective probability measure. Recall that the rows of the memory matrix M correspond to elements of context, while the columns correspond to item features. Hence, each element of M represents the frequency with which

[20] One might think, then, that $\zeta = 1$ is "optimal" for decision-makers who value faster response times; however, there may be a trade-off because high ζ also means low autocorrelation of context during encoding, which is problematic if decision-making depends on understanding contiguity of events. Thus, the "optimal" ζ is likely strictly between 0 and 1.

[21] Unlike ζ, however, η also governs the degrees of mutual inhibition and decision variance, so that its total effect on the decision process is much more complicated than just multiplying the response time.

[22] Wachter and Kahana (in press) give this same interpretation to their context-based encoding process. In that paper, context is interpreted as an unobservable feature, not a neural state. The math is the same.

a given item has been observed with a given context. Let us define $\hat{M}$ as the memory matrix M rescaled so that all of its elements sum to one:

$$\hat{M}_t = \frac{M_t}{\iota_m^\top M_t \iota_n},$$

where ι_k is a k-vector of ones. It follows that each element of $\hat{M}$ represents the historical, unconditional joint probability of observing a given item in a given context: $\hat{M}_t(j,i) = p_t(\tilde{c}_j, f_i)$, where $\tilde{c}_j$ denotes the jth m-dimensional standard basis vector. Under the additional assumption that elements of context represent unconditional probabilities over context elements (i.e., $c_t(j) = p_t(\tilde{c}_j)$), we show that the context-based model implies Bayesian learning.

5.2.1 Encoding as Bayesian learning

Suppose features y_i are presented and encoded in increasing order. Upon the perception of feature y_i the subject retrieves input context c_i^{in} as the ith column of the normalized matrix $M_{i-1}(\Gamma_{i-1}^{\text{col}})^{-1}$ (recall (1)). Because $M_{i-1}(\Gamma_{i-1}^{\text{col}})^{-1}$ scales the columns of M_{i-1} by their total frequency, Bayes' rule implies that the jth element of its ith column equals $p_i(\tilde{c}_j|f_i)$, the conditional probability of the jth context given the observation of item i. Context then updates as a weighted average of input context and past context, such that

$$c_i(j) = p_i(\tilde{c}_j) = (1-\zeta)p_{i-1}(\tilde{c}_j) + \zeta p_i(\tilde{c}_j|f_i).$$

This is of the same form as standard temporal-difference reinforcement learning rules.[23] This context is then interacted with features and stored back into memory as

$$\hat{M}_i = (1-\theta_i)\hat{M}_{i-1} + \theta_i c_i f_i^\top,$$

where the weights

$$\theta_i = \frac{\phi_i}{\iota_m^\top M_0 \iota_n + \sum_{i'=1}^{i} \phi_{i'}}.$$

[23] In the language of the reinforcement-learning literature, we may think of $p_{i-1}(\tilde{c}_j) - p_i(\tilde{c}_j|f_i)$ as a prediction error and of ζ as a learning-rate parameter.

This is another temporal-difference learning rule, except that now the learning rate declines over time.[24] Notice that the ith column of the matrix $c_i f_i^\top$ simply equals $p_i(\tilde{c}_j)$. This means that the joint probabilities stored in $\hat{M}$ become

$$\hat{M}_i(j, i') = p_i(\tilde{c}_j, f_{i'}) = \begin{cases} (1-\theta_i)p_{i-1}(\tilde{c}_j, f_{i'}) + \theta_i p_i(\tilde{c}_j) & \text{if } i' = i, \\ (1-\theta_i)p_{i-1}(\tilde{c}_j, f_{i'}) & \text{if } i' \neq i. \end{cases}$$

The unconditional probability of observing the current item increases, and this increase is distributed across context rows in proportion to the current context vector.

5.2.2 Retrieval as Bayesian inference

We begin by shutting down competitive retrieval ($\eta = 1$). During the retrieval and response phase, the subject begins by retrieving a set of features f_0^{in} from the initial context c_0 and the normalized matrix $M_0^\top(\Gamma_0^{\text{row}})^{-1}$. By Bayes' rule, the columns of $M_0^\top(\Gamma_0^{\text{row}})^{-1}$ store the conditional probabilities $p(f_i|\tilde{c}_j)$. Therefore, again using Bayes' rule, retrieved features equal

$$f_t^{\text{in}}(i) = \sum_{j=1}^{m} p(f_i|\tilde{c}_j)p_t(\tilde{c}_j) = p_t(f_i).$$

That is, f^{in} is simply a vector of individual item probabilities.

We now turn to the role of iterative retrieval. Recalling the definition of Φ from (11), some algebra reveals that

$$\Phi(i, i') = \sum_{k=1}^{m} p(f_i|\tilde{c}_k)p(\tilde{c}_k|f_{i'}).$$

In words, elements of Φ map features f to contexts c (through $p(c|f)$), then map contexts back to features (through $p(f|c)$) in probability space. To see this more clearly, consider the updating rule for Δf^{in}, which we can write out as

$$f_t^{\text{in}} = (1-\zeta)f_{t-1}^{\text{in}} + \zeta\Phi f_{t-1}^{\text{in}}.$$

[24] Again comparing this to reinforcement learning, the prediction error is now $c_i f_i^\top - M_{i-1}$ and the learning rate is θ_i. This parameter is a decreasing function of time, so this is a constant-gain learning rule.

At the beginning of period t, the subject has a set of time-$(t-1)$ probability assessments stored in f_{t-1}^{in}. Let us denote these by $f_{t-1}^{\text{in}}(i) = p_{t-1|t-1}(f_i)$ to emphasize that these probabilities were formed at $t-1$ using contemporaneous information. The subject then uses Φ to extract new information from these probabilities via Bayes' rule: the ith element of the product Φf_{t-1}^{in} is

$$\begin{aligned}[\Phi f_{t-1}^{\text{in}}](i) &= \sum_{i'=1}^{n}\sum_{k=1}^{m} p(f_i|\tilde{c}_k)p(\tilde{c}_k|f_{i'})p_{t-1|t-1}(f_{i'}) \\ &= \sum_{k=1}^{m} p(f_i|\tilde{c}_k)p_{t|t-1}(\tilde{c}_k) \\ &= p_{t|t-1}(f_i).\end{aligned}$$

The notation $p_{t|t-1}$ emphasizes that this is an updated probability at time t using information from the previous period. Finally, the subject weights this updated probability with the previous assessment to get a fully updated probability

$$f_t^{\text{in}}(i) = p_{t|t}(f_i) = (1-\zeta)p_{t-1|t-1}(f_i) + \zeta p_{t|t-1}(f_i).$$

Yet again, we have a temporal-difference learning rule, this time over item probabilities.

When we reintroduce competitive retrieval, probabilities $\tilde{p}_{t|t}(i) = \tilde{f}_t^{\text{in}}(i)$, calculated as above, undergo the nonlinear transformation F to generate decision probabilities

$$p_{t|t}(i) = \frac{\tilde{p}_{t|t}(i)^{\eta}}{\sum_{i'=1}^{n}\tilde{p}_{t|t}(i')^{\eta}}.$$

While this kind of transformation is not a standard element of Bayesian learning, it is a very common element of the statistical learning literature, going back as far as Bush and Mosteller (1955) and Luce (1959). This kind of transformation is also commonplace in the reinforcement-learning literature, which assigns choice probabilities via a convex transformation of perceived action values (e.g., a softmax function).

6 Concluding remarks

Although theories of memory have come to account for varied aspects of memory encoding and retrieval, these theories have not offered a mechanistic account of how the memory system retrieves encoded information and maps it to the environment. Various retrieval rules that have been offered merely assume that, somehow, the memory system retrieves the most similar target among a sea of competitors. Here we propose a solution to this problem within the setting of retrieved-context theories of memory retrieval.

The idea behind our model is simple: the same processes that underlie the encoding of information into memory also underlie retrieval. Individuals use memory to call experiences to mind, and those recalled experiences serve as a jumping-off point for the recollection of more experiences. People follow a train of thought until it leads to the recovery of information that is useful for the task at hand.

We formalize this notion of following a train of thought via two new mechanisms. The first is an *iterative retrieval* process, whereby experiences are sequentially sampled from memory. The presence of a latent mental context serves an important role in this sequential sampling, because it facilitates the mapping between features of the environment and representations in the mind. Context retrieves features, then features retrieve context, and so on.

The second core mechanism underlying our model is a *competitive retrieval* process. Under competitive retrieval, those features that come to mind most prominently suppress the recollection of less prominent features. Competitive retrieval serves two purposes. First, it allows individuals to map retrieved information back to recognizable features from the environment. Second, it ensures that the features that do ultimately come to mind are related to the circumstances at the time of retrieval. This is especially important in decision-making situations, because most decisions require context-dependent information.

To validate our framework, we apply the model to the standard setting of a free-recall experiment in which a subject studies a series of items and subsequently attempts to recall as many as possible in any order. Simulating a reduced-form version of retrieved-context theory that includes our iterative retrieval model produces the classic effects of recency, primacy, and temporal contiguity. Of particular interest in the present study, we examined

whether the model could also produce the fat-tailed distribution of inter-response times seen in the data and the faster inter-response times seen when transitioning between neighboring list items. This is indeed the case, suggesting that our framework can provide an accurate quantitative account of the retrieval and response process.

Our framework sheds light on the role of memory in evidence-accumulation models of retrieval and decisions (e.g., Brown & Heathcote, 2008; Ratcliff, 1978; Usher & McClelland, 2001). In diffusion-based retrieval models, some initially retrieved information serves as the input to a noisy accumulation process that evolves to a boundary. We show how the key mechanisms in diffusion-based models can be mapped back to our two novel mechanisms — iterative and competitive retrieval — suggesting that our framework can explain how the memory system underlies these popular accumulation processes. Notably, our model features no exogenous noise in the retrieval process and therefore provides a new explanation for the large variation in choices and response times observed in many settings.

By better understanding how the memory system is involved in retrieval and response times, we believe that our framework makes an important stride in reconciling principles of memory with principles of decision-making. The encoding and retrieval of information no doubt plays a central role in how individuals make choices. Our hope is that the present framework can serve as a stepping stone for future work pursuing the much loftier goal of understanding how memory shapes decisions.

7 Acknowledgements

We thank Sudeep Bhatia, Greg Cox, David Halpern, Benjamin Hébert, Marc Howard, Andrei Shleifer, Jessica Wachter, and conference and seminar participants at the Context and Episodic Memory Symposium, the Society for Mathematical Psychology Conference, and the University of Pennsylvania for helpful comments.

APPENDIX

A Derivations for Section 3

A.1 Derivation of retrieved-feature dynamics

We first translate the expressions from Section 2 to their continuous-time analogues. Note that we can substitute the expression (1) into (2) and rewrite it in terms of changes:

$$\Delta c_t = \zeta(M_0(\Gamma_0^{\text{col}})^{-1} f_{t-1}^{\text{in}} - c_{t-1}). \tag{A.1}$$

The continuous-time equivalent of this context evolution is:

$$\frac{dc_t}{dt} = \zeta(M_0(\Gamma_0^{\text{col}})^{-1} f_t^{\text{in}} - c_t). \tag{A.2}$$

To get the evolution of f_t^{in}, we need only take a time derivative; this is feasible due to the assumption that the competitive-retrieval function F is differentiable at all of its elements. Each element of f_t^{in} is (potentially) a function of all the elements of $\tilde{f}_t^{\text{in}}$. Thus, by the chain rule, the derivative of the ith element of f_t^{in} with respect to time t is

$$\frac{df_t^{\text{in}}(i)}{dt} = \sum_{i'=1}^{n} \frac{df_t^{\text{in}}(i)}{d\tilde{f}_t^{\text{in}}(i')} \frac{d\tilde{f}_t^{\text{in}}(i')}{dt}. \tag{A.3}$$

Since $\tilde{f}_t^{\text{in}} = M_0^\top (\Gamma_0^{\text{row}})^{-1} c_t$, the ith element of $\tilde{f}_t^{\text{in}}$ is

$$\tilde{f}_t^{\text{in}}(i) = m_0^{\text{row}}(i) c_t, \tag{A.4}$$

where $m_0^{\text{row}}(i)$ is the ith row of $M_0^\top (\Gamma_0^{\text{row}})^{-1}$. This means that

$$\frac{df_t^{\text{in}}(i)}{dt} = \sum_{i'=1}^{n} \frac{\partial f_t^{\text{in}}(i)}{\partial \tilde{f}_t^{\text{in}}(i')} m_0^{\text{row}}(i') \frac{dc_t}{dt} \tag{A.5}$$

Substituting (A.2) into this expression, we have

$$\frac{df_t^{\text{in}}(i)}{dt} = \sum_{i'=1}^{n} \frac{\partial \tilde{f}_t^{\text{in}}(i)}{\partial \tilde{f}_t^{\text{in}}(i')} \zeta(\Phi^{\text{row}}(i') f_t^{\text{in}} - \tilde{f}_t^{\text{in}}(i')), \tag{A.6}$$

where $\Phi^{\text{row}}(i)$ is the ith row of the matrix Φ, as defined in equation (11). In matrix notation, this means that the full system evolves according to

$$\frac{df_t^{\text{in}}}{dt} = \zeta D^F(\tilde{f}_t^{\text{in}})(\Phi f_t^{\text{in}} - \tilde{f}_t^{\text{in}}), \tag{A.7}$$

where D^F is an $n \times n$ Jacobian matrix with elements

$$D^F(i, i') = \frac{\partial f_t^{\text{in}}(i)}{\partial \tilde{f}_t^{\text{in}}(i')}. \tag{A.8}$$

Adding and subtracting $\zeta D^F(\tilde{f}_t^{\text{in}}) f_t^{\text{in}} dt$ to the right-hand side of (A.7), we have

$$\frac{df_t^{\text{in}}}{dt} = \zeta D^F(\tilde{f}_t^{\text{in}})(\Phi - I) f_t^{\text{in}} + \zeta D^F(\tilde{f}_t^{\text{in}})(f_t^{\text{in}} - \tilde{f}_t^{\text{in}}), \tag{A.9}$$

the continuous-time version of the discrete-time process (13) stated in the main text.

A.2 Derivation for power rule

If the competitive-retrieval rule F takes a power form (9), then

$$f_t^{\text{in}}(i) = \frac{\tilde{f}_t^{\text{in}}(i)^\eta}{\sum_{i'=1}^{n} \tilde{f}_t^{\text{in}}(i')^\eta} \tag{A.10}$$

In this case, the derivatives of D^F are

$$D^F(i, i') = \begin{cases} -\eta \dfrac{\tilde{f}_t^{\text{in}}(i)^\eta \tilde{f}_t^{\text{in}}(i')^{\eta-1}}{(\sum_{i'=1}^{n} \tilde{f}_t^{\text{in}}(i')^\eta)^2} & \text{if} \quad i \neq i', \\ \eta \dfrac{\tilde{f}_t^{\text{in}}(i)^{\eta-1}}{\sum_{i'=1}^{n} \tilde{f}_t^{\text{in}}(i')^\eta} - \eta \dfrac{\tilde{f}_t^{\text{in}}(i)^{2\eta-1}}{(\sum_{i'=1}^{n} \tilde{f}_t^{\text{in}}(i')^\eta)^2} & \text{if} \quad i = i'. \end{cases} \tag{A.11}$$

Now, using the definition of f^{in}, this simplifies to

$$D^F(i,i') = \begin{cases} -\eta f_t^{\text{in}}(i) f_t^{\text{in}}(i') \tilde{f}_t^{\text{in}}(i')^{-1} & \text{if} \quad i \neq i', \\ \eta f_t^{\text{in}}(i) \tilde{f}_t^{\text{in}}(i)^{-1} - \eta f_t^{\text{in}}(i)^2 \tilde{f}_t^{\text{in}}(i)^{-1} & \text{if} \quad i = i'. \end{cases} \tag{A.12}$$

Collecting terms,

$$D^F(i,i') = \begin{cases} -\eta f_t^{\text{in}}(i)(f_t^{\text{in}}(i')/\tilde{f}_t^{\text{in}}(i')) \leq 0 & \text{if} \quad i \neq i', \\ \eta(1 - f_t^{\text{in}}(i))(f_t^{\text{in}}(i)/\tilde{f}_t^{\text{in}}(i)) \geq 0 & \text{if} \quad i = i'. \end{cases} \tag{A.13}$$

We can re-express this in matrix notation as

$$D^F(\tilde{f}_t^{\text{in}}) = \eta(I - f_t^{\text{in}} \iota_n^\top) D^{\text{ratio}}(\tilde{f}_t^{\text{in}}) \tag{A.14}$$

where ι_n is an n-dimensional column vector of ones and $D^{\text{ratio}}(\tilde{f}_t^{\text{in}})$ is an $n \times n$ diagonal matrix with ith diagonal

$$D^{\text{ratio}}(i,i) = \frac{f_t^{\text{in}}(i)}{\tilde{f}_t^{\text{in}}(i)}. \tag{A.15}$$

B Proof of convergence in iterative retrieval

Recall (15) from the main text: absent competitive retrieval (i.e., for $\eta = 1$), retrieved features at time t equal

$$f_t^{\text{in}} = A^t f_0^{\text{in}}, \quad A = (1-\zeta)I + \zeta\Phi.$$

Because, as we explain in the main text, Φ has one eigenvalue equal to one and all other eigenvalues less than one in magnitude, it has an eigenvalue decomposition

$$\Phi = U\Lambda U^\top,$$

where U is an orthonormal matrix of eigenvectors and Λ is a diagonal matrix of eigenvalues:

$$\Lambda(i,i) = \begin{cases} 1 & \text{if} \quad i = 1, \\ \lambda_i \in (-1,1) & \text{if} \quad i \neq 1. \end{cases}$$

Using the fact that $UU^\top = I$, it follows that

$$A = (1-\zeta)I + \zeta U \Lambda U^\top = U \tilde{\Lambda} U^\top,$$

where

$$\tilde{\Lambda} = (1-\zeta)I + \zeta\Lambda = \begin{cases} 1 & \text{if} \quad i = 1, \\ \tilde{\lambda}_i = (1-\zeta) + \zeta\lambda_i & \text{if} \quad i \neq 1. \end{cases}$$

That is, A also has an eigenvalue decomposition with one eigenvalue equal to one and all others $|\tilde{\lambda}_i| < 1$, and with corresponding eigenvectors identical to those of Φ. In particular, this means that $\lim_{t\to\infty} \tilde{\Lambda}^t = \lim_{t\to\infty} \Lambda^t$, and so

$$f_\infty^{\text{in}} = \lim_{t\to\infty} A^t f_0^{\text{in}} = \lim_{t\to\infty} U\tilde{\Lambda}^t U^\top f_0^{\text{in}} = \lim_{t\to\infty} U\Lambda^t U^\top f_0^{\text{in}} = \lim_{t\to\infty} \Phi^t f_0^{\text{in}}.$$

In other words, f_∞^{in} is also the right eigenvector of Φ corresponding to its eigenvalue of one, and is hence the unique vector satisfying

$$f_\infty^{\text{in}} = \Phi f_\infty^{\text{in}},$$

as claimed in the main text.

References

Aka, A., Phan, T. D., & Kahana, M. J. (2021). Predicting recall of words and lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(5), 765–784. http://dx.doi.org/10.1037/xlm0000964

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765. https://doi.org/10.1037/0033-295x.113.4.700

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice reaction time: Linear ballistic accumulation. *Cognitive Psychology*, *57*, 153–178.

Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. John Wiley & Sons, Inc.

Greene, N. R., Goldman, S. T., & Kahana, M. J. (2024). Inter-response times in free recall. *Manuscript in preparation*.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299. https://doi.org/10.1006/jmps.2001.1388

Kahana, M. J. (2012). *Foundations of human memory*. New York, NY: Oxford University Press.

Kahana, M. J. (2020). Computational models of memory search. *Annual Review of Psychology*, *71*(1), 107–138. 10.1146/annurev-psych-010418-103358

Kahana, M. J., Aggarwal, E. V., & Phan, T. D. (2018). The variability puzzle in human memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(12), 1857–1863. https://doi.org/10.1037/xlm0000553

Kahana, M. J., Lohnas, L. J., Healey, M. K., Aka, A., Broitman, A. W., Crutchley, P., ... Weidemann, C. T. (2024). The Penn electrophysiology of encoding and retrieval study. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/xlm0001319

Long, N. M., & Kuhl, B. A. (2019). Decoding the tradeoff between encoding and retrieval to predict memory for overlapping events. *NeuroImage*, *201*, 116001. https://doi.org/10.1016/j.neuroimage.2019.07.014

Luce, R. D. (1959). *Individual choice behavior*. John Wiley.

Masson, M. E. J. (1995). A distributed memory model of semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 3–23. https://doi.org/10.1037/0278-7393.21.1.3

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, *116*(1), 129–156. https://doi.org/10.1037/a0014420

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59-108.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review*, *9*(3), 438–481. 10.3758/BF03196302

Rizzuto, D. S., & Kahana, M. J. (2001). An autoassociative neural network model of paired-associate learning. *Neural Computation*, *13*(9), 2075–2092. https://doi.org/10.1162/089976601750399317

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review*, *115*(4), 893–912. https://doi.org/10.1037/a0013396

Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, *90*(5), 927–39. https://doi.org/10.1016/j.neuron.2016.04.036

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*(3), 550–592. https://doi.org/10.1037/0033-295x.108.3.550

Wachter, J. A., & Kahana, M. J. (in press). A retrieved-context theory of financial decisions. *Quarterly Journal of Economics*. 10.3386/w26200