Bridging CRU and CMR in Free and Serial Recall: A Factorial Comparison of Retrieved-Context Models

Jordan B Gunn and Sean M Polyn
Department of Psychology, Vanderbilt University

Author Note

The authors are grateful to Gordan Logan and Adam Osth for their constructive feedback on earlier versions of this manuscript.

Abstract

Retrieved-context theory posits that episodic retrieval is driven by a representation that evolves over time, tying each item to the contextual features present during encoding and later accessing those associations to guide retrieval (Howard & Kahana, 2002). Although the Context Maintenance and Retrieval (CMR) model (Morton & Polyn, 2016; Polyn et al., 2009) offers a flexible and well-tested retrieved-context implementation – addressing both free and serial recall (Lohnas, 2024) – it is relatively complex, incorporating mechanisms absent in some other retrieved-context models. In contrast, the Context Retrieval and Updating (CRU) model (Logan, 2018, 2021) provides a simpler, more streamlined specification of context-driven retrieval shown to excel in strictly ordered memory tasks like serial recall. However, it remains unclear whether CRU's leaner architecture extends easily to unconstrained retrieval dynamics in free recall. It is similarly unknown whether CMR's added mechanisms confer meaningful advantages over leaner CRU in serial recall. To investigate the gap between CRU and CMR, we systematically compare them, presenting them side by side and exploring how each can be viewed as a parameterized variant of the same foundational ideas. Using a factorial model selection approach, we selectively incorporate CMR-like features into CRU and compare each hybrid variant to standard CMR on free and serial recall data. We find that selectively incorporating CMR-like features substantially improves CRU's fit to free recall and that CRU's item-confusion and recall termination mechanisms can help CMR capture serial recall data.

Keywords: episodic memory, free recall, serial recall, retrieved-context theory, computational modeling

Bridging CRU and CMR in Free and Serial Recall: A Factorial Comparison of Retrieved-Context Models

Introduction

Researchers have long recognized that episodic memory does not involve merely storing items but also associating them with the temporal context in which they occur. Early perspectives on temporal context (Estes, 1955; Melton & Martin, 1972; Tulving & Madigan, 1970) proposed that a slowly drifting set of features provides a unique background signal at each point in time, helping the mind keep track of when certain items were experienced. Retrieved-context theory (Howard & Kahana, 1999, 2002) extends these ideas further, suggesting that when an item is recalled, it reactivates the contextual state that was present during its original encoding, thus cueing neighboring items from the same temporal window. Because this context representation reflects a gradually shifting set of features, reactivating a given item's context can bring to mind items from adjacent positions.

The Temporal Context Model (TCM; Howard & Kahana, 2002) crystallized these ideas into a formal model, simulating how a drifting context vector could explain robust patterns in free recall such as the recency effect (in which items from the end of a list are often recalled first) and the temporal contiguity effect (in which items studied near one another in time tend to be recalled successively). Subsequent work showed that TCM could be enriched to capture additional empirical findings, giving rise to a range of models that extend the basic retrieved-context principles to account for a variety of memory phenomena (Howard & Kahana, 2002; Kahana, 2020; Lohnas et al., 2015; Sederberg et al., 2008). Among these extensions, the Context Maintenance and Retrieval (CMR) model (Morton & Polyn, 2016; Polyn et al., 2009) built upon TCM's evolving context representation by adding flexible associative mechanisms, including the incorporation of pre-experimental structures and semantic relationships. The model has been applied to tasks ranging from free recall to recognition memory and has explained a broad array of behavioral and neural data (Healey & Kahana, 2016; Horwath et al., 2023; Kragel et al., 2015; Morton & Polyn, 2016; e.g., Polyn et al., 2009; Sederberg et al., 2011).

Against this background, Gordon Logan has extended the retrieved-context perspective to explain how automatic serial order behavior can emerge in tasks like typing or immediate serial recall (Logan, 2018, 2021). Logan's work on automaticity (Logan, 1988) and hierarchical control underscores that once a top-level goal (for example, "type DOG") is set, a series of lower-level actions can be generated with minimal conscious oversight – an idea that is realized in the Context Retrieval and Updating (CRU) model. CRU assumes that each keystroke or list item is associated with a continuously updated context state and that retrieving one item integrates that item into context, thereby cueing retrieval of the next. After initially demonstrating CRU's potential as an account of automatic typing (Logan, 2018), CRU has been applied to capture a wide range of serial recall phenomena, including transposition gradients, serial position effects, and error patterns (Logan, 2021), with even further work clarifying its contributions against models focused on positional representations of serial order (Logan, 2021; Logan & Cox, 2023; Osth & Hurlstone, 2023). Importantly, CRU highlights item identification mechanisms that handle confusable stimuli (e.g., visually similar letters) and adopts a streamlined approach to context updating. By design, it aspires to offer a general retrieved-context account of sequential performance in domains like typing, musical performance, and language production.

Although CMR was designed primarily for free recall, recent work by Lohnas (2024) has shown that a variant called sCMR can also capture key features of serial recall, hinting at a unified framework spanning both free and serial tasks. That analysis highlights assumptions in CRU that may limit its ability to address free recall, such as its lack of pre-experimental context-to-feature associations and its simplified context updating scheme. Although CRU was proposed as a general model of sequence retrieval, it has so far been tested mainly in serial recall tasks with visually confusable stimuli, leaving open the question of whether its streamlined architecture can address performance in free-recall tasks where participants can recall items in any order. We aim to address this gap by systematically comparing CRU and CMR, exploring how each model can be viewed as a parameterized variant of the same foundational ideas, and clarify whether CRU's simpler approach remains domain-bound or can be extended to capture free recall phenomena traditionally captured by CMR. In parallel, we extend the same factorial logic to serial recall, asking whether CMR's added machinery delivers gains when strict order is required – or merely adds complexity.

In our approach, we analyze CRU and CMR as two ends of a parameterized spectrum with broadly overlapping assumptions. Starting from the Morton and Polyn (2016) CMR baseline and the Logan (2021) CRU baseline, we fit both with the same event-by-event likelihood that scores each recall in sequence. Each model treats recall as context-driven, linking items to an evolving temporal context that then guides what is retrieved next, but they differ in how they handle item confusions, associative structures, recall initiation and termination, and differences in encoding across study positions. We treat these differences as modular components that can be selectively enabled in a factorial manner, creating a family of hybrid variants that systematically isolates the impact of each mechanism. With this predictive framework, we fit these variants to empirical data from word-based free recall (Healey & Kahana, 2014) and letter-based serial recall (Logan, 2021), testing how each mechanism contributes to capturing basic behavioral patterns in these tasks. By identifying which model-specific assumptions boost performance, we aim to clarify how CRU and CMR can be integrated into a more unified account of episodic memory search, highlighting the shared principles that underlie both models and the unique mechanisms that each brings to the table.

Model Structure

Overall, CRU is a streamlined version of CMR with fewer parameters and a simpler architecture for updating context, though it adds some processing steps absent in CMR. CMR employs dual feature-to-context and context-to-feature linear associative memory matrices to mediate the evolution of context and context-driven retrieval, respectively. By contrast, CRU pairs items directly with context states in an instance-based memory and omits an explicit feature-to-context memory, relying on a simplified context updating scheme suitable for serial recall. However, CRU also incorporates specialized mechanisms absent from CMR, including an item-identification stage that resolves visually confusable inputs (e.g., *p* vs. *q*) and a response-selection stage that allows output confusions; the latter helps explain articulatory errors in serial recall (Osth & Hurlstone, 2023). In this study we test the item-identification mechanism only for serial-recall fits and disable both stages for the word-based free-recall data, where such confusions are negligible and challenging to model outside the scope of letter-based stimuli.

We outline each model's key features in turn, focusing on how they (i) represent and encode items, (ii) initialize and evolve context, (iii) store item–context associations, and (iv) handle recall—especially stopping rules and transitions. Throughout, we note convergences and divergences, emphasizing that CRU and CMR share a family resemblance in how context guides retrieval, while still offering distinct approaches to certain subtasks. These differences may at times be interpreted as competing explanations of the same phenomena or as complementary extensions useful in different domains of memory research. Across compatible features, we can view each parameter configuration as a strategy chosen to meet tasks demands (Logan & Gordon, 2001; Zhang et al., 2023): baseline CRU could represent a less flexible forward-chaining strategy optimized for strict serial order, whereas baseline CMR could represent a broader retrieval strategy that can be adapted to free recall. The factorial hybrids we create therefore span a continuum of alternative yet compatible retrieved-context strategies that different tasks or different individuals can recruit.

In later sections, we examine whether selectively adding certain "missing" CMR-like mechanisms (e.g., dynamic feature-to-context learning or pre-experimental associations) can help CRU capture free recall's hallmark backward transitions and flexible stopping. We also explore whether CRU's item-identification mechanism can help CMR capture the strict order of serial recall, and whether CMR's additional mechanisms contribute meaningfully to serial recall performance.

Item Identification in CRU

CRU and CMR each represent items as orthonormal unit vectors. However, CRU can optionally model confusability among similar items via a probabilistic competition step at encoding. This helps the model account for patterns of intrusion and other errors in serial recall tasks, where participants may confuse similar items (e.g., p vs. q). CRU can also model confusion errors that occur during retrieval, but following the reference CRU implementation described by Logan (2021), we do not include this mechanism in our present evaluation. When item i is presented, a racing diffusion with drift rates scaled by g determines which representation crosses the threshold first:

$$v_i = \exp\left[-g \cdot d_{ij}\right].$$

Here, v_i is the drift rate for item i, d_{ij} is the distance between the presented item and each candidate j, and g is a sensitivity parameter. Lower g fosters more confusion among similar items, whereas higher g makes identification more accurate.

Because confusion errors receive limited treatment in the study of word-based free recall and are challenging to model outside the scope of letter-based stimuli, we disable the item identification mechanism in our free recall fits (i.e., set $g \rightarrow \infty$). In contrast to CRU, CMR has historically assumed perfect item recognition for tasks like word recall and uses no item identification mechanism. In our evaluation of letter-based serial recall, we evaluate CRU's item identification mechanism as a possible extension of CMR.

Context Initialization and Evolution: Equivalent Between Models

Both CMR and CRU represent context as a vector of continuous values. They initialize this vector by setting one dimension to 1.0 and the rest to 0.0. In both models, this initial state represents the start-of-list context and can be reinstated to organize retrieval.

This context vector evolves as items are encoded and retrieved, integrating a new contextual input at each step. At each step i, both models update context as

$$\mathbf{c_i} = \rho_i \mathbf{c_{i-1}} + \beta\, \mathbf{c_i^{\mathrm{IN}}},$$

where β controls integration of new input $\mathbf{c_i^{IN}}$, and ρ_i normalizes the vector:

$$\rho_i = \sqrt{1 + \beta^2\left[(\mathbf{c_{i-1}} \cdot \mathbf{c_i^{IN}})^2 - 1\right]} \; - \; \beta(\mathbf{c_{i-1}} \cdot \mathbf{c_i^{IN}}).$$

This gradual integration yields a recency-based gradient reflecting the order in which items were presented.

Contextual Input: CRU Omits Feature-To-Context Learning

A key difference between CRU and CMR is how each model obtains $\mathbf{c_i^{IN}}$ from the active item. CRU simply treats the one-hot (orthonormal) item representation as $\mathbf{c_i^{IN}}$. Since each unique item representation is orthogonal to all others, this direct assignment ensures that the context vector is updated in a way that is unique to the item being studied or recalled. By contrast, CMR uses a feature-to-context memory M^{FC} that learns experimental associations between items and context states when items are presented.

Like CRU, CMR's M^{FC} is initialized by associating each item representation with a unique context unit such that when an item is presented, the corresponding context unit is activated and integrated into the context vector as contextual input. But unlike CRU, in CMR a learning rate parameter γ configures the relative influence of pre-experimental and experimental associations in this memory. Pre-experimental associations capture these baseline or default item–to-context connections according to:

$$M^{FC}_{\text{pre}(ij)} = \begin{cases} 1-\gamma, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases}$$

Thus, M^{FC}_{pre} resembles a partial identity matrix, ensuring that each item i is initially tied to a corresponding context unit i with weight $1-\gamma$. Experimental associations are then formed during encoding whenever an item is presented, based on the simultaneous activity of item and context features f_i and c_j, where γ serves as the learning rate:

$$\Delta M^{FC}_{ij} = \gamma f_i c_j$$

This update rule is a Hebbian outer product, capturing the principle that co-active item features and context states become associated. Accordingly, M^{FC} serves as a matrix of Hebbian associations that stores both the baseline connections (scaled by $1 - \gamma$) and any newly formed links (scaled by γ).

In CMR, when an item is recalled, these pre-experimental and experimental associations allow the model to reinstate the contextual state that was present when the item was originally presented. This retrieved context contains fading traces of contextual information associated with preceding items. Once this contextual information (stored in M^{FC}) is retrieved, it can support backwards transitions in free recall. By contrast, CRU's direct assignment of $\mathbf{c_i^{IN}}$ corresponds to CMR with a learning rate $\gamma = 0$. This causes encoding or retrieval of an item to always integrate an orthogonalized context vector unique to that item, eliminating a source of support for backward transitions in sequence recall.

The absence of a feature-to-context memory in CRU (vs CMR) reflects how the temporal contiguity effect – the tendency to transition between items that were presented close together in time (Kahana, 1996, 2020) – manifests differently in serial and free recall tasks. Since participants in serial recall are typically asked to reproduce the studied list in order, the temporal contiguity effect is almost entirely forward-going, with participants transitioning to the next item in the list. But since free recall participants are free to recall items in any order, the temporal contiguity effect is more clearly bidirectional, with participants often transitioning to the item that was presented immediately before the retrieved item. A dynamic feature-to-context memory like CMR's may be more effective in capturing this bidirectional effect, while CRU's simpler approach may suffice for serial recall.

Context-to-Item Associations: CRU Omits a Pre-Experimental Structure

CMR uses a context-to-feature matrix M^{CF}, seeded with pre-experimental associations via parameters δ (self-support) and α (shared support):

$$M^{CF}_{\text{pre}(ij)} = \begin{cases} \delta, & \text{if } i = j \\ \alpha, & \text{if } i \neq j \end{cases}$$

Here,δ works like γ in M^{FC} to scale the pre-experimental context-to-feature associations compared to the experimental associations acquired during the experiment. By contrast, the parameter α approximates uniform semantic support between all items. When a context unit is activated, any nonzero α provides partial activation to all items, providing a coarse way to capture semantic clustering effects without simulating the semantic identity of individual study items (Polyn et al., 2009). Some versions of CMR replace this uniform structure with richer semantic representations (Morton & Polyn, 2016; Sederberg et al., 2008).

CRU, by contrast, lacks explicit pre-experimental context-to-item associations, effectively setting $\delta = 0$ and $\alpha = 0$. In CMR, increasing δ while γ is 0 tends to exclusively heighten short-lag backward transitions in free recall, similar to how raising γ (the feature-to-context learning rate) promotes backward transitions. When γ is freed, δ's role is broader, increasing the likelihood of short-lag transitions in free recall compared to long-lag, wide-ranging transitions. Because most serial recall tasks emphasize forward order and rarely show backward transitions, CRU omits pre-experimental context-to-item associations in keeping with its streamlined approach.

It is worth noting that TCM's original formulation (Howard & Kahana, 2002) also did not include a pre-experimental context-to-feature matrix, focusing instead on how a drifting context signal promotes both recent and temporally adjacent items during recall. Subsequent extensions – e.g., TCM-A (Sederberg et al., 2008) – introduced pre-experimental context-to-feature associations to capture the recency effect displayed by patients with amnesia, along with other parameters and retrieval decision rules to address short- vs. long-term recency. CMR (Morton & Polyn, 2016; Polyn et al., 2009) incorporated explicit pre-experimental structures like δ and α to capture semantic clustering and backward transitions more effectively.

During study, CMR updates M^{CF} via a Hebbian rule on linear associative connections, whereas CRU encodes item–context pairs as separate memory traces in an instance-based framework. Despite these architectural differences between linear associative networks and instance-based models, the two approaches can sometimes exhibit functionally similar dynamics (J. R. Anderson & Bower, 1972; Turner, 2019). For example, a linear associator (as in CMR) can approximate an instance-based model (as in CRU) when the input patterns are sufficiently distinct, effectively storing and retrieving each studied episode with minimal interference. We do not explore these equivalences in detail here, but we treat both implementations of context-to-item associations as serving a broadly similar role: binding items to context so that recalling one item can cue related items. Evaluating tasks with repeated items or strong semantic structure where these distinctions become more critical lies outside the scope of this paper but remains a compelling direction for future research.

Serial Position Effects

Both serial and free recall tasks exhibit primacy and recency effects, where participants exhibit better memory for items from the beginning and end of study lists, respectively (Murdock, 1962). A separate but related observation is that participants often initiate recall with either the first or the last item in a list. CRU and CMR handle primacy and recency effects differently, reflecting differences in their theoretical emphases and the tasks they are designed to address.

In CMR, two parameters ϕ_{s} and ϕ_{d} configure the learning rate of its context-to-feature memory M^{CF} to enforce a primacy gradient, scaling the amount of learning based on the serial position of the studied item according to:

$$\phi_i = \phi_{\mathrm{s}} e^{-\phi_{\mathrm{d}}(i-1)} + 1$$

Here, ϕ_i is the learning rate for the i-th item, ϕ_{s} is the learning rate at the first serial position, ϕ_{d} is the decay rate, and i is the serial position of the studied item. The strengths of M^{CF} associations is thus updated according to:

$$\Delta M_{ij}^{CF} = \phi_i c_j f_i$$

By contrast, CRU modulates its sensitivity parameter g and context-integration rate β by serial position. This is done by setting a maximum value for each parameter and a decay rate that scales the maximum value according to the serial position of the studied item such that the integration rate and sensitivity at the i-th serial position are:

$$g_i = g_{\text{max}} \cdot g_{\text{dec}}^{i-1}$$

$$\beta_i = \beta_{\text{max}} \cdot \beta_{\text{dec}}^{i-1}$$

Here, g_{max} and β_{max} are the initial values of g and β, and g_{dec} and β_{dec} are the decay rates. The equations for identifying items and updating context are left unchanged, except using the position-specific values of these parameters.

Along with modulating memorability as a function of serial position, the CRU and CMR models also address serial position effects by configuring the state of context when retrieval begins. In CMR, the end-of-list contextual state can be modified by retrieving and integrating the start-of-list contextual state according to a special integration rate parameter β_{start} before initiating retrieval (Healey & Wahlheim, 2024). When this integration rate is zero, the contextual cue primarily targets items near the end of the study list, and when it is high, the cue targets items near the beginning of the study list. Hence CMR can flexibly shift recall initiation to either early or late list items, whereas CRU always begins from the start-of-list context.

This mechanism is not present in CRU. Instead, retrieval always begins with the start-of-list contextual state, ensuring that initial items in the study list are always most accessible. This is equivalent to setting $\beta_{\text{start}} = 1$ in CMR. The differences in how CRU and CMR account for serial position effects in memory search in part reflect differing emphases in the models. CMR was primarily designed to address free recall, where both primacy and recency effects are prominent (Murdock, 1962), motivating a β_{start} parameter that can be adjusted to target different parts of the study list. By contrast, in serial recall tasks, primacy effects are more salient than recency effects, and the start-of-list contextual state is always the most relevant cue for retrieval. CRU's simpler approach to context initialization and retrieval reflects this emphasis on serial recall, where the goal is to reproduce the studied list in order.

Recall and Recall Termination

In either model, retrieval attempts compare the current context to all stored contexts in the context-to-feature memory, prioritizing items associated with contexts that are most similar to the current context. CRU performs this comparison directly to each context stored in its instance-based memory using a dot product. CMR's context-to-feature memory performs this comparison by projecting the current context through its context-to-feature matrix M^{CF} to activate item representations. In either model, this context-to-item comparison sets up a probabilistic competition between items for retrieval where items that are more strongly associated with the current context are more likely to be retrieved. Both models also allow the possibility of terminating recall without retrieving any more items, though they differ in how they factor this possibility into the competition between items. In either model, once a retrieval attempt is complete, the item is reported, the context is updated based on the retrieved item, and the process repeats until recall is terminated. The most complete specification of CRU also includes a mechanism where retrieved items can be incorrectly reported due to typos or other errors. We omit this mechanism here (as in its original evaluation, Logan (2021)), but other work has shown that this mechanism can be useful in capturing articulatory errors in serial recall (Osth & Hurlstone, 2023).

CRU treats the end of the study list as a special "end-of-list" event, assigning it a representation just like any study item and associating it with the final contextual state. Because this representation resides in the same instance-based memory as all other contexts, it competes with studied items during retrieval: if the end-of-list representation wins, recall terminates. As a result, CRU predicts that termination is most likely immediately after recalling the final item and least likely after the first item.

Different implementations of CMR treat recall termination differently. In several implementations of CMR, a leaky, competitive accumulation process determines which item wins each recall competition (Healey & Kahana, 2016; Lohnas et al., 2015; e.g., Polyn et al., 2009). In these implementations, each simulated recall takes a certain amount of time, and recall terminates when the elapsed recall time exceeds the duration of the recall period. It is computationally expensive to estimate the likelihood of any given recall event using the leaky, competitive accumulation process, leading to the development of a simplified recall termination process amenable to likelihood-based model evaluation (Hong et al., 2024; Kragel et al., 2015; Morton & Polyn, 2016). With this simplified framework, CMR treats recall termination as a separate process from item retrieval that does not depend on the content of the current context or on which items have been recalled so far. For CMR, the probability of stopping the recall sequence without retrieving any more items, depends on the output position j and is given by:

$$P(\text{stop}, j) = \theta_{\text{s}} e^{j\theta_{\text{r}}}$$

Here, θ_{s} and θ_{r} govern the initial stopping probability and the rate at which this probability increases, respectively, which is modeled as an exponential function.

CRU and CMR also differ in how they traditionally model the choice between items during retrieval. CMR applies a probabilistic choice rule (Luce, 1959) to convert activation strengths into recall probabilities. An additional parameter τ controls the contrast in activation strengths between items by raising the activation of each item to the power of τ. The probability $P(i)$ of recalling a specific item i is thus:

$$P(i) = (1 - P(\text{stop}))\frac{A_i^{\tau}}{\sum_k^N A_k^{\tau}}$$

In contrast, CRU models the choice between options as race between independent processes where each runner (here, a contextual state stored in memory) is governed by a drift rate parameter v and a threshold θ. This threshold is normally set to a fixed constant, allowing the rate parameter to exclusively determine the winner. The rate parameter for each option is determined by the similarity between the current context and the context associated with the item. The finishing time is characterized by the Wald distribution:

$$f(t) = \theta(2\pi t^3)^{-1/2} \cdot exp(-\frac{(\theta - t \cdot v)^2}{2t})$$

And the race between runners is characterized by the following, with the actual probability that the runner *i* wins being the integral of this function over time:

$$f(t,i) = f_i(t) \prod_{j \neq i}^{N} (1 - F_j(t))$$

The way either model approaches the choice between items during retrieval is not normally treated as core to their theoretical commitments. Instead, the choice of decision rule is typically justified based on computational tractability and simplicity, with the probabilistic choice rule being more common in CMR and the racing diffusion model being more common in CRU. Furthermore, the racing diffusion model allows for the possibility of modeling response times in addition to accuracy, though implementations of CRU in serial recall have not yet leveraged this feature. Given that neither model is committed to a specific decision rule, we will prefer to use matched decision rules in our model comparisons to ensure that differences in model performance are not due to differences in decision rules. Runner drift rates configured for simulation of the racing diffusion process can be treated as item support values in application of the probabilistic choice rule, and vice versa, to ensure that the models are compared on a level playing field. This leaves evaluation of the relative merits of the racing diffusion and probabilistic choice rules for modeling memory search within a retrieved-context framework as a topic for future work.

Overall, CMR and CRU differ in how flexibly they address the generation of retrieval sequences. First, CMR includes two separate parameters to model recall termination, θ_{s} and θ_{r}, which control the initial tendency to stop and the rate at which this tendency grows over successive recalls. Second, CMR includes a sensitivity parameter τ that controls the contrast in activation strengths between competing items. This parameter can be evaluated independently of the probabilistic choice rule for its ability to capture recall performance, similarly to the sensitivity parameter g in CRU's item identification mechanism. In addition to these already mentioned differences, CMR uses separate integration rate parameters for encoding and retrieval (β_{enc} and β_{rec}) to control how much the context integrates with each new item during encoding and retrieval.

Evaluation Approach

While CRU was originally developed to capture behavioral performance in tasks focused on serial order, many of CRU's structural differences from CMR do not necessarily hinge on serial ordering. We therefore ask which of these differences matter for free recall, where retrieval can proceed in any order and generally features more frequent backward transitions and flexible stopping. Conversely, given that CMR has been shown able to address serial recall (Lohnas, 2024), we also ask whether CMR's added mechanisms confer any advantages over CRU in capturing the strict order of serial recall, or whether CRU's specialized mechanisms for addressing item confusability can provide further improvements to CMR's performance in this domain. To do so, we adopt a factorial approach in which each key mechanism that CRU omits (or implements differently than CMR) is selectively toggled on or off, creating a series of "hybrid" CRU variants that, in a sense, span the gap between CMR and CRU in a hypothetical space of all possible models. Our evaluation is not comprehensive – we do not explore every possible combination of toggled mechanisms – but instead focuses on the most salient differences between CRU and CMR. In particular, we focus on five key factors that distinguish CRU from CMR: (1) dynamic feature-to-context memory, (2) pre-experimental context-to-feature associations, (3) serial position scaling, (4) recall initiation, (5) recall termination, and (6) confusable item identification.

Factors and Variants

In order to permit a fair comparison between CRU and CMR, we equate the two models in several ways. First, some parameters are allowed to vary freely across all model variants (e.g., τ, β_{enc}, and β_{rec}) to avoid complicating the factorial design further. Second, the item identification factor is omitted in free recall fits, as it is challenging and unusual to address outside the scope of letter-based stimuli. Instead, examination of this factor is limited to the serial recall fits. Finally, we use the same decision rule (the probabilistic choice rule) for all model variants, ensuring that differences in model performance do not arise from different decision processes. The resulting framework generates model variants that differ along the following dimensions:

1. **Dynamic or Inert Feature-to-Context Memory**. In CMR, a nonzero learning rate γ

allows recently presented items to retrieve the context states present during the study period, enhancing backward transitions in free recall. By contrast, CRU omits M^{FC} (or sets $\gamma = 0$), instead directly integrating item representations into the context vector with no reinstatement of previously associated contexts. This factor is especially pertinent in free recall, where backward transitions are a key empirical signature. We show that enabling a dynamic M^{FC} in CRU can help capture these transitions and improve its overall fit to free recall data. In serial recall, this factor is less relevant, but may still help capture phenomena like the fill-in/in-fill effect, where omissions are more likely to be followed by the preceding item than the following item (Logan & Cox, 2023; Lohnas, 2024).

2. **Use or Nonuse of a Pre-Experimental Context-to-Feature Memory**. Both α (shared support) and δ (self-support) in CMR reflect pre-existing item associations that can facilitate retrieval. CRU does not include these mechanisms. We consolidate these into a single on/off factor indicating whether such pre-experimental support is active. When on, items have baseline associations that may enhance recall probabilities – particularly relevant in word-based free recall, where semantic or lexical relationships matter. When off, all associations arise solely from the experimental context, mirroring CRU's instance-based memory. By toggling this factor, we show that incorporating pre-experimental support can help CRU capture patterns of temporally distant and adjacent transitions in free recall and can potentially help CRU address patterns of backward transitions in serial recall.

3. **Primacy Gradient**. A third factor involves how each model boosts memory for initial study items to account for primacy effects. CRU achieves this by modulating contextual integration rate β with β_{dec} and item identification sensitivity g by serial position, causing the strength of these mechanisms to peak at the start of the list and decay toward the end. CMR instead employs dedicated parameters ϕ_s and ϕ_d that scale the learning rate by serial position in context-to-feature memory M^{CF} from an initial peak to a final baseline. We don't evaluate item confusability in free recall ($g \rightarrow \infty$), so for free recall evaluation we

focus on assessing the effectiveness of CMR-style primacy gradient in CRU. We show that incorporating CMR's primacy gradient can enhance CRU's ability to capture the primacy effect in free recall without supposing that items can be visually confused with one another during encoding. In our serial recall evaluation, we have the item confusability mechanism and variable contextual integration rate activated across variants, and observe that incrementally adding CMR's primacy gradient has limited impact on performance.

4. **Recall Initiation**. CRU always initiates recall from the start-of-list context, whereas CMR can bias the cue toward the beginning or end of the list via β_{start} by integrating the final context state back toward the start. We toggle this factor by either forcing retrieval to begin at the start-of-list state (CRU style) or allowing a flexible mix of start and end contexts (CMR style). In free recall, participants frequently begin recalling from the end of the list. We show that allowing CRU to flexibly initiate retrieval is crucial for its ability to address this pattern and free recall performance more generally, but that such flexibility is not as relevant for addressing the strict order of serial recall.

5. **Recall Termination**. CRU encodes an "end-of-list" state that competes with items to terminate recall, whereas CMR separately calculates a stopping probability using θ_{s} and θ_{r}. In our design, we can adopt either the CRU-style item-based mechanism or the CMR-style exponential stop rule. This factor addresses how each model handles the decision to stop retrieval – especially in free recall, where participants spontaneously cease recall rather than exhaustively listing items in order. Here we show that CRU's context-based termination mechanism causes the model to collapse when addressing free recall datasets with strong recency effects, and that incorporating a CMR-style stopping probability can help CRU better capture these patterns. However, by contrast, we show that CRU's end-of-list mechanism is more effective than CMR's stopping probability in capturing the strict order of serial recall. It's possible that in serial recall participants explicitly encode and retrieve an end-of-list marker as a signal that typing is done (Logan, 2018), but that in

free recall participants instead attempt to continue recalling items until it is too difficult to do so or time runs out.

6. **Confusable Item Identification**. CRU's item identification mechanism can be toggled on or off, allowing it to model confusability among similar items during encoding and retrieval. When freeing both g and g_{dec}, the mechanism also helps address serial position effects by scaling the likelihood of confusion errors by serial position. Evaluation of this factor is hindered in word-list free recall by the lack of empirical confusability statistics for the items in the PEERS dataset compared to the letter-based stimuli in the serial recall dataset. Furthermore, evaluation of this factor is complicated even in our serial recall fits by the fact that no alternative mechanism for predicting item intrusions is available in CMR for comparison. We leave CRU's g and g_{dec} parameters free across all model variants to enable prediction of item intrusions. Our evaluation examines whether the mechanism can be readily incorporated into CMR to help it capture patterns of intrusion and other errors in serial recall tasks and whether additionally toggling on CMR's mechanisms for addressing serial position effects can further improve model performance. We reproduce the results of Logan (2021) to confirm that CRU's item identification mechanism can help capture patterns of intrusion and other errors in serial recall tasks across study and output positions.

By toggling or switching between mechanisms – e.g., enabling or disabling feature-to-context learning – we create a spectrum of CRU variants that bridge between CRU's streamlined architecture and CMR's more flexible associative structures. We also include a "baseline CMR" and a "baseline CRU" variant that represents the full CMR and CRU models, respectively, to compare against the hybrid variants.

Datasets

We draw our free recall data from the PEERS dataset (Healey & Kahana, 2014), focusing on a subset well-suited to our model comparisons. This dataset has already been used to evaluate CMR in various contexts (Healey & Kahana, 2014; Morton & Polyn, 2016), and we use it here to evaluate CRU’s performance in free recall. **Participants.** A total of 126 individuals completed multiple trials where they were instructed to study lists of words and recall them in any order. **Stimuli.** Each list consisted of 16 words drawn at random from the Toronto Word Pool (Friendly et al., 1982), which includes common nouns, adjectives, and verbs. Words were chosen to be low in semantic similarity to minimize confusability; thus, items generally had negligible overlap. **Procedure.** Participants studied each list of words, then performed immediate free recall for up to 45 seconds. They repeated this process for multiple lists, although trials featuring an interleaved retrieval task were excluded from our analysis to maintain consistency with standard free recall protocols. In total, this exclusion left 112 trials per participant. **Rationale.** We selected this dataset because (a) it provides a large, diverse sample size (allowing robust parameter estimation), (b) it uses word stimuli, typical for CMR applications, and (c) it exhibits well-documented primacy and recency effects, presenting a challenge for CRU’s traditional emphasis on strictly serial tasks.

We draw our serial recall data from the serial recall trials of Experiment 1 by Logan (2021), which was collected to evaluate CRU’s performance in a range of serial report tasks. **Participants.** A total of 24 participants completed multiple trials where they were presented with lists of letters and instructed to report them in their order of order. We examine the subset of serial recall trials where participants began typing the letters once the list was removed. **Stimuli**. Each list consisted of 5, 6, or 7 letters drawn at random from the English alphabet. **Procedure**. In serial recall trials, participants were presented with the letters simultaneously on a computer screen for 1000 ms, and then the letters were removed from the screen before the participants typed their response. After each trial, a 1000-ms blank screen was presented before the next trial began. **Rationale.** We selected this dataset because (a) it provides many trials per participant (allowing robust parameter estimation), (b) it uses letter stimuli, enabling clear evaluation of CRU’s item identification mechanism, and (c) it exhibits common patterns of serial recall, setting up a clear comparison between CRU and CMR.

Likelihood-Based Fitting and Simulation

We fit each model variant at the individual-participant level using a step-wise likelihood procedure common in recent CMR work and embraced in CRU's 2021 evaluation (Kragel et al., 2015; Logan, 2021; Morton & Polyn, 2016; Polyn, 2023). For every trial the model (1) encodes the study list, updating context and associative matrices, then (2) scores each recall event in sequence, adding the log-probability of the observed outcome and updating the model's state before the next event.

Log-likelihoods are summed across trials and maximised with differential evolution (Storn & Price, 1997). Since differential evolution is a stochastic optimization algorithm, we run it 10 times for each participant and model variant, retaining the best-performing fit. To compare models, we report mean and bootstrapped 95% confidence intervals for the log-likelihood scores across participants for each variant. We also examine which variant gives the best fit on a participant-by-participant basis, probing the robustness of our findings and whether there are individual participants for whom CRU or CMR is a better fit despite the overall group-level results.

In evaluation of free recall data, consistent with prior CMR evaluations, repeated recalls and extralist intrusions are filtered from trials before the likelihood calculation. By contrast, for serial recall data, repetitions and intrusions are considered theoretically diagnostic and thus are not filtered out. Instead, we allow models evaluated with serial recall data to make both repetition and intrusion errors.

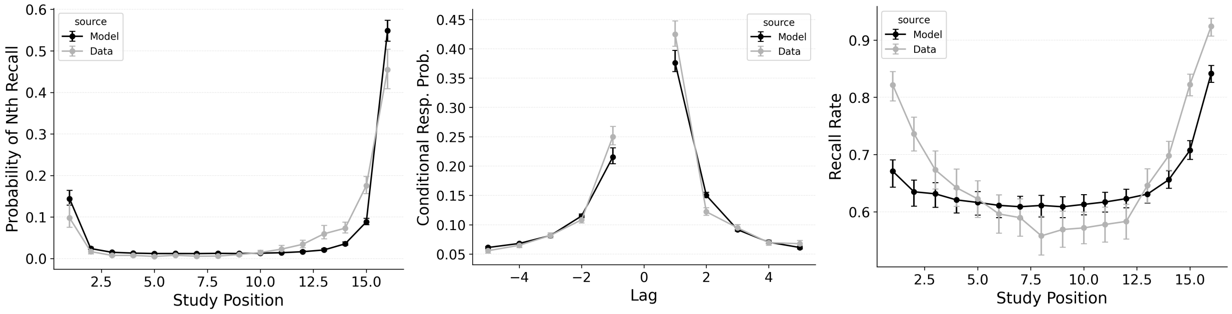
CRU's item confusability mechanism presents special challenges for likelihood-based fitting, as model state is not necessarily fully determined by observed study and recall events. In this case, the score for an observed recall event is not simply the probability of retrieving the observed item, but a combination of the probability of retrieving the observed item after a successful encoding event, and the probability of reporting the item as a result of confused/failed encoding (i.e., a visual-identification slip, in Logan's terminology). The probability of typing letter j equals the sum, over all studied items i, of that item's retrieval probability times its chance to slip into j. After scoring a recall, context is advanced with the maximum-a-posteriori (MAP) retrieved item – the one that maximises the product $P_{\text{mem}}(i)P_{\text{slip}}(j \mid i)$. Using this approach across evaluations of applicable model variants ensures they are compared on a level playing field even when the output includes repetitions and intrusion errors.

We also simulate each fitted model variant on the same list structure participants experienced to see whether it reproduces hallmark free and serial recall phenomena. By comparing real and simulated data, we assess whether a model not only matches the trial-by-trial recall sequences but also generates the correct shape of these benchmark patterns. Since adding additional parameters almost always improves fits, we use the ability of model variants to simulate these patterns as our primary metric of model performance.

Free Recall Results

Figure 1

Summary statistic fits of the baseline CMR model to PEERS data. ***Left****: probability of recall initiation by serial position.* ***Middle****: conditional response probability as a function of lag.* ***Right****: recall probability by serial position.*



Our first evaluation uses the Healey and Kahana (2014) subset of the PEERS free recall dataset. Using the overall goodness of fit for each model variant, standard CMR provides the best fit at the group level (see Table 2) as well as for 100% of individual participants. This is no surprise, as CMR uses more parameters and was designed to capture free recall phenomena while CRU was developed for serial recall tasks. Our focus instead is in demonstrating how specific model mechanisms improve CRU’s fit to the benchmark behavioral phenomena of free recall: serial position effects, recall initiation effects, and temporal organizational effects, as depicted in

Figure 1.

A serial position analysis Figure 1 demonstrates both a primacy effect (a memorability advantage for early-list items) and a recency effect (a memorability advantage for late-list items). These memorability advantages are also apparent in the distribution of first recalls Figure 1. In this dataset, a recency effect is particularly pronounced in the distribution of first recalls; participants usually start recall with the last item from a study list, but sometimes start with the first item. By comparison, the serial position curve shows a more balanced trade-off between primacy and recency effects.

Two mechanisms allow CMR to capture the interplay between primacy and recency effects. First, CMR leverages the associative primacy gradient mechanism modulating learning rates in the context-to-feature memory (ϕ_s, ϕ_d) to enhance the memorability of early-list items. Second, CMR recovers the start-of-list context with a dedicated context integration parameter (β_start). Both mechanisms influence the initiation of recall and the overall serial position curve, but they have differential effects on these two stages. The β_start parameter is most influential at recall initiation, while the associative primacy gradient influences the primacy effect throughout retrieval. These mechanisms together allow CMR to capture both the distribution of first recalls and the overall serial position curve.

A focal concern of retrieved-context models is the lag-contiguity effect, the tendency to successively recall items that were presented close together in time (Kahana, 1996, 2020). These models capture this effect by reinstating the temporal context associated with a just-recalled item (step 1) and then using that updated context to cue other items (step 2). In free recall, the lag-contiguity effect is bidirectional but asymmetric, favoring forward transitions over backward ones. To capture all aspects of the lag-contiguity effect, CMR primarily uses its feature-to-context memory to reinstate the context associated with a just-recalled item and then its context-to-feature memory to activate items associated with its newly updated context. The γ parameter influences context reinstatement by weighting the influence of pre-experimental and experimental associations in the feature-to-context memory. CMR's parameters δ and α conversely configure pre-experimental context-to-feature memory to influence the extent to which retrieval of items associated with the current context is focused on nearby or distant neighbors of the last recalled item. Together, these parameters yield the characteristic forward-biased yet bidirectional pattern observed in free recall.

CMR exhibits some limitations capturing these phenomena, underpredicting the primacy effect in the serial position curve. CMR's ability to account for performance using this dataset has been evaluated before (Healey & Kahana, 2014; Morton & Polyn, 2016); depending on whether CMR is fit to summary statistics or to individual response sequences, and on how many rerolls of the differential evolution fitting algorithm are used, CMR's exact fit to these phenomena can vary. Recent work suggests that a mixture model, in which a given trial either has support for the end-of-list items or support for the start-of-list items, may better capture the dynamics of recall initiation (Osth & Farrell, 2019). Addressing these limitations and further exploring these interactions are outside the scope of this paper and reflect continuing challenges in modeling free recall phenomena. These simulation results demonstrate the ability of standard CMR to capture key free-recall phenomena, establishing a baseline that can be built upon in future work. We now turn to CRU and evaluate how its streamlined implementation impacts its ability to capture these phenomena.

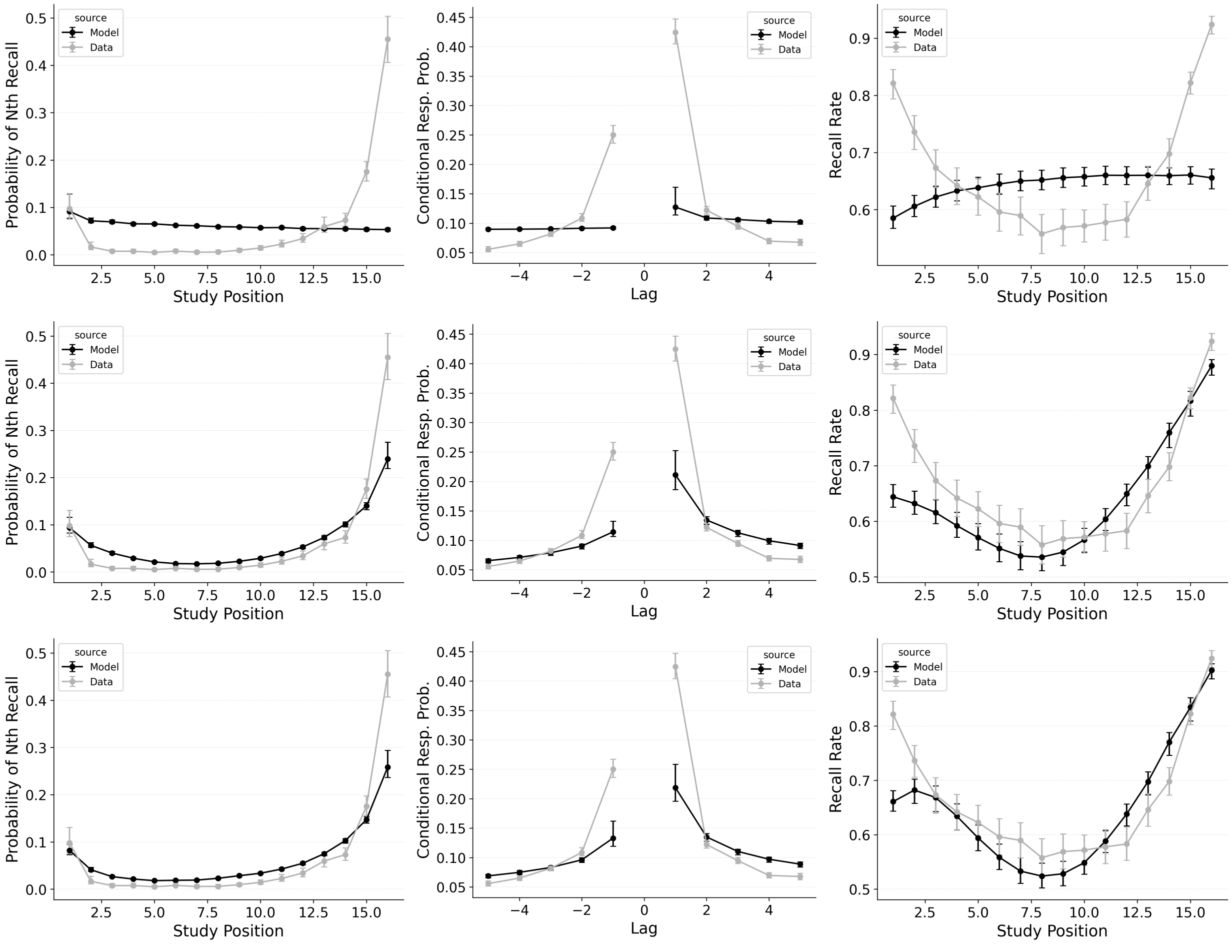
Addressing Serial Position Effects in Recall Initiation and Overall

CRU includes most of CMR's core mechanisms. However, CRU's default start-of-list recall initiation mechanism forces it to strongly prioritize the start of the list during recall initiation, which causes its performance to collapse when fit to free recall data exhibiting a strong recency effect. This failure to capture recall initiation affects the entire response sequence because transitions in free recall depend substantially on prior recalls. Allowing CRU to initiate retrieval with a blend of the end-of-list and start-of-list context according to a flexible start-of-recall parameter like CMR's β_{start} substantially improves its ability to capture these phenomena Figure 2.

Figure 3 shows the impact of shifting the start context integration rate parameter β_{start} on the probability of starting recall by serial position and the recall probability by serial position for

Figure 2

*Summary statistic fits of baseline CRU (**Top**), CRU with free start context integration rate β_{start} (**Middle**), and CRU freeing both start context integration rate (β_{start}) and primacy gradient (ϕ_s and ϕ_d) parameters (**Bottom**) to PEERS free recall data. **Left**: probability of recall initiation by serial position. **Middle**: conditional response probability as a function of lag. **Right**: recall probability by serial position.*



CMR. To perform this simulation, we fit CMR to each individual participant in the subset of the PEERS dataset (Healey & Kahana, 2014) and then shifted the β_{start} parameter from 0 to 1 in increments of 0.1, repeatedly simulating the model on the same list structure it was fit to and generating serial position curves and recall initiation curves. This parameter primarily trades off the strength of the primacy effect in recall initiation against the strength of the recency effect, with higher values of β_{start} leading to stronger primacy effects and lower values leading to stronger recency effects. The serial position curve is sensitive to the value of β_{start} as well since the item retrieval initiates with affects the trajectory of responses throughout the recall sequence.

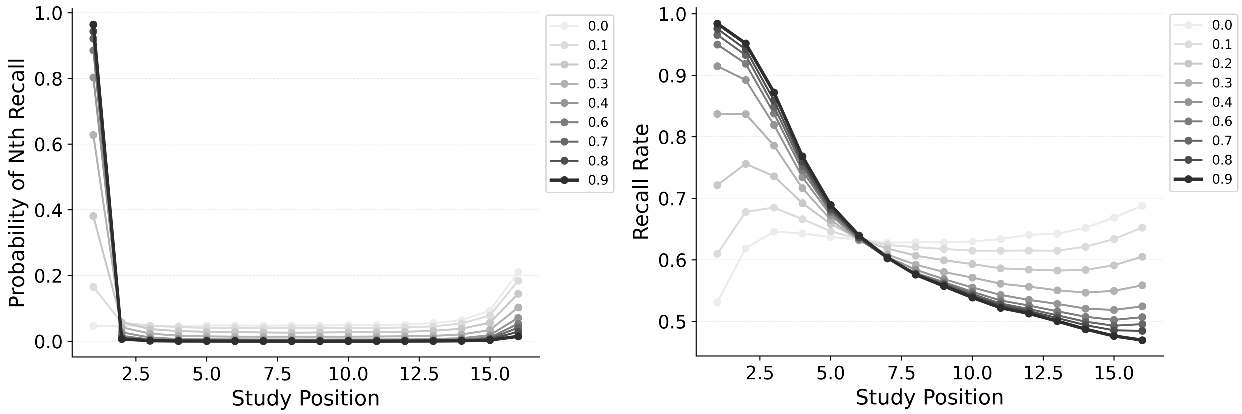
We illustrate this point by comparing standard CRU to a CRU variant that includes CMR's start context integration rate β_{start} parameter as well as a CRU variant that additionally includes CMR's primacy gradient (ϕ_{s} and ϕ_{d}) parameters. Rows 1 and 2 of Figure 2 show simulated benchmark phenomena for CRU and the first variant, respectively. Fixing β_{start} to 1 leads to fits where CRU predicts no consistent serial position or lag-contiguity effects, while allowing β_{start} to vary enables CRU to begin to capture these effects, achieving a U-shaped overall serial position curve, a strong recency effect in recall initiation, and a bidirectional lag-contiguity effect. The model still substantially underestimates the primacy effect in the overall serial position curve, the strength of the recency effect in recall initiation, and the strength of the lag-contiguity effect, but these limitations are less severe than when β_{start} is fixed at 1.0.

Further allowing CRU to include CMR's primacy learning gradient (ϕ_{s} and ϕ_{d} parameters) to modulate context-to-feature memory learning rates to peak at the start of the study list further improves its performance. This variant is better at capturing the strength of the primacy effect in the overall serial position curve Figure 2, but does not as substantially improve the model's ability to capture the recency effect in recall initiation or the lag-contiguity effect.

These differences observed between CRU and CMR in their ability to address serial position effects may be exaggerated by the exclusion of an item identification mechanism in CRU addressing free recall. Logan (2021) accounted for primacy effects in CRU with g_{max} and g_{dec} parameters that allowed the model to modulate the sensitivity of item identification by serial

Figure 3

*Simulation of the impact of shifting the start context integration rate parameter β_{start} on the probability of starting recall by serial position (**Left**) and the recall probability by serial position (**Right**) for CMR. Using parameters fit to PEERS free recall data, β_{start} is shifted from 0 to 1 in increments of 0.1, with the color of the lines indicating the value of the parameter.*



position, but this mechanism does not apply to word free recall data. Here, only CRU's $\beta_{\max}$ and β_{dec} parameters were available to modulate contextual integration rates during encoding as a function of serial position, which may have limited the model's ability to capture the primacy effect in the overall serial position curve. In our forthcoming evaluation of the models using serial recall data, the full CRU model is examined.

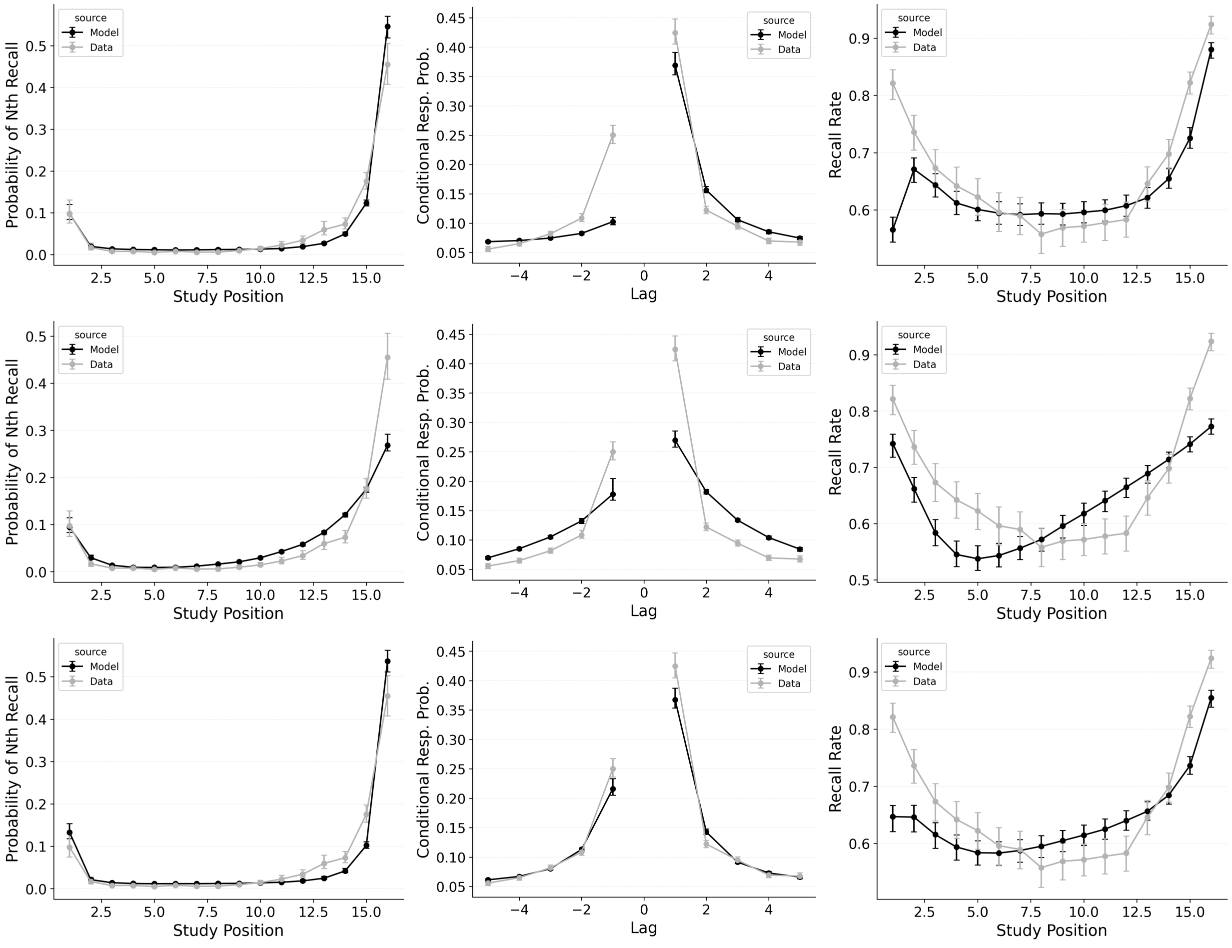
Tuning the Sharpness and Asymmetry of the Lag-Contiguity Effect

In free recall, the lag-contiguity effect is bidirectional and asymmetric, with participants more likely to transition to items that were presented immediately after the just-recalled item, but also sometimes transitioning to items that were presented immediately before the just-recalled item in the original study list. Standard CRU is able to simulate a bidirectional lag-contiguity effect, but it substantially underestimates the strength of the effect compared to the data, predicting around a 20% probability of +1 lag transition while the data shows over a 40% probability of +1 lag transition. Standard CRU shows a similar discrepancy for a backward -1 lag transition.

CMR controls the shape of the lag-contiguity effect with a set of parameters controlling

Figure 4

Summary statistic fits of models to the PEERS free recall dataset (Healey & Kahana, 2014). ***Top****: CRU with free pre-experimental context-to-feature memory (α, δ), primacy gradient (ϕ_s, ϕ_d), and start context integration rate (β_{start}) parameters.* ***Middle****: CRU with free item-to-context learning rate (γ), primacy gradient (ϕ_s, ϕ_d), and start context integration rate (β_{start}) parameters.* ***Bottom****: CRU with free item-to-context learning rate (γ), pre-experimental context-to-feature memory (α, δ), primacy gradient (ϕ_s, ϕ_d), and start context integration rate (β_{start}) parameters – equivalent to CMR.* ***Left****: Probability of starting recall by serial position.* ***Left****: Probability of starting recall by serial position.* ***Middle****: Conditional response probability as a function of lag.* ***Right****: Recall probability by serial position.*



the relative strength of different associative structures. These parameters include γ, which influences context reinstatement during retrieval, and α and δ, which influence the competition between items during retrieval. For larger values of the γ parameter, the context associated with a just-recalled item is more stronglyreinstated during retrieval, helping capture high rates of short lag backward transitions in the lag-contiguity effect. Figure 5 (left) shows how altering γ affects the lag-CRP for CMR, supporting this interpretation. The simulation was configured similarly to those described in the previous section, using parameters fit to Healey and Kahana (2014), and shifting γ from 0 to .9 in increments of 0.1.

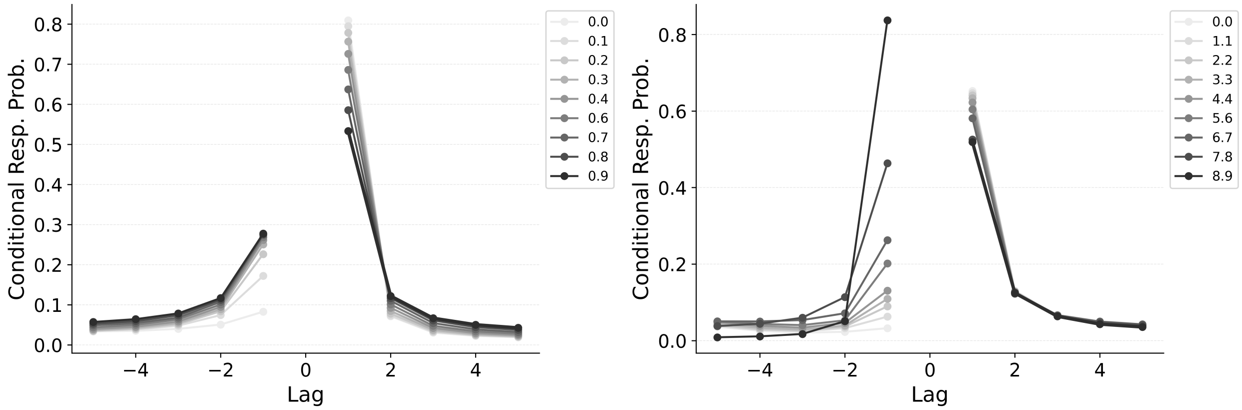
CMR's δ and α parameters provide a pre-experimental context-to-feature memory that can influence the lag-contiguity effect by tuning the strength of associations from context representations to item representations. When δ is much higher than α, the lag-contiguity effect is facilitated by favoring neighbors of the just-recalled item in the competition between items during retrieval. Alternatively, setting δ to match α or to zero can flatten the lag-contiguity effect by making all items equally likely to be activated by a context feature associated with a just-recalled item. When γ is configured to 0 – as it effectively is in CRU – tuning δ no longer influences rates of backward transitions in the lag-contiguity effect, as the experimental associations that would facilitate these backward transitions are no longer present. Figure 5 (right) illustrates the impact of shifting δ on the shape of the lag-CRP for CMR when γ is not set to 0. The simulation was configured similarly to the previous section, using parameters fit to Healey and Kahana (2014), and shifting δ from 0 to 8.9 in increments of 1. When γ is configured to non-zero values, higher values of δ can simultaneously drive both forward and backward transitions in the lag-contiguity effect, reducing the rate of more distant transitions while increasing the rate of closer transitions.

Freeing γ, δ, and α in CMR allows the model to flexibly capture the strength and asymmetry of the lag-contiguity effect. Figure 4 (Row 2) illustrates CRU's performance when extended to include CMR's dynamic feature-to-context memory (γ) alongside primacy and recency mechanisms (ϕ_s, ϕ_d, and β_{start}). Simulated summary statistics confirm that this extension allows the model to better capture the lag-contiguity effect, but the model still underestimates its overall strength compared to the data. This produces correspondingly worse performance capturing primacy and recency benchmarks.

Finally, in the third row of Figure 4, extending CRU to include both CMR's dynamic feature-to-context memory (γ) and pre-experimental context-to-feature memory (δ and α) alongside established primacy and recency mechanisms allows the model to better capture all components of the lag-contiguity effect and other benchmarks. However, this model is exactly CMR: CRU with all of CMR's mechanisms enabled. This suggests that standard CRU's streamlined implementation is not sufficient to capture the full range of free recall phenomena, and underlines that all of CMR's mechanisms are useful for capturing free recall data, despite the complexity they introduce.

Figure 5

*Simulation of the impact of shifting CMR's γ (**Left**) and δ (**Right**) parameters on the conditional response probability as a function of lag for CMR. Using parameters fit to Healey and Kahana (2014), the learning rate parameter γ is shifted from 0 to 1 in increments of 0.1, and the item support parameter δ is shifted from 0 to 10 in increments of 1, with the color of the lines indicating the value of the parameter.*

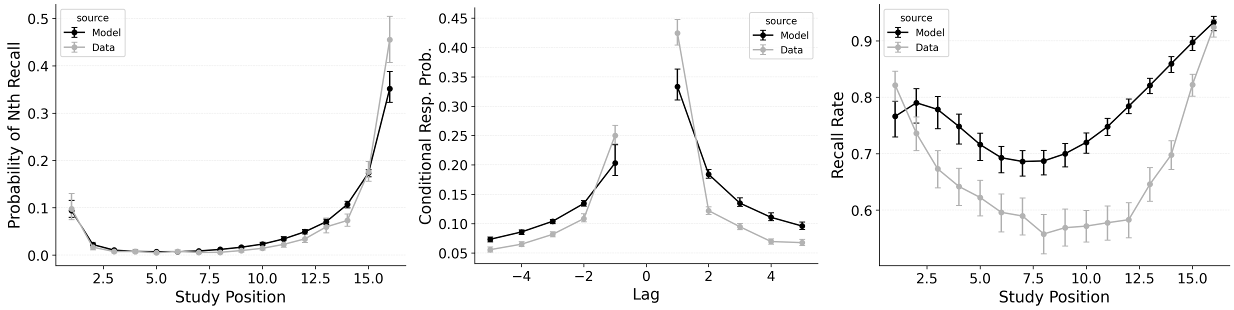


Position- vs Context-Based Mechanisms for Recall Termination

CRU and CMR's mechanisms for recall termination are fundamentally different. This difference cannot be captured by toggling a parameter from a fixed to a freely adjustable value,

Figure 6

Summary statistic fits of CMR with CRU's context-based recall termination mechanism to Healey and Kahana (2014). ***Left****: probability of starting recall by serial position.* ***Middle****: conditional response probability as a function of lag.* ***Right****: recall probability by serial position.*



but rather, these mechanisms can only be swapped between variants. CMR uses exponentially increasing stopping probabilities θ_s and θ_r to model recall termination; the probability of termination scales only with the number of recalls made so far. By contrast, CRU treats the end of a study sequence as a special item associated in memory with the final state of the study context. This item competes with other items for retrieval at each new recall event, and its activation can terminate recall. In this specification, the probability of termination depends on the state of context at each recall event, and can be influenced by the same mechanisms that influence the probability of recalling other items.

Performance differences between CMR's position-based recall termination mechanism and CRU's context-based recall termination mechanism are substantial. While Figure 1 shows baseline CMR's performance on these benchmarks, Figure 6 shows the performance of CMR with CRU's context-based recall termination mechanism. While patterns in response initiation are well-captured, using context-based recall termination mechanism leads CMR to predict overly high recall rates for all study list positions and to fail to capture the sharpness of the lag-contiguity effect. The success of CRU's context-based recall termination mechanism depends on how consistently participants terminate recall after recalling the final items from the study list. In most serial recall datasets, participants tend to perform this way, and the mechanism correspondingly predicts that the probability of terminating recall scales with the number of recalls made so far, as context drifts from its start-of-list state to its end-of-list state. By contrast, in free recall datasets where participants exhibit a strong recency effect in recall initiation, this mechanism can predict early termination of recall upon or even before retrieving the last item in the study list.

Serial Recall Results

Our goal is to determine whether any of the mechanisms that give CMR an advantage in free recall can improve the fit of CRU to this serial recall dataset (Expt. 1 from Logan (2021), described above). Importantly, we allow item identification confusability via CRU's g and g_{dec} parameters to be freely estimated in all model variants so that intrusion errors can be predicted. Without this mechanism the likelihood of intrusions is 0. Given the presence of intrusions in the data, this collapses the fitting process.

The overall goodness of fit for each model variant is presented in Table 3. The models are ordered by overall goodness of fit. These log-likelihood differences are highly reliable at the group level, based on wAIC comparisons (not reported). The baseline version of CRU is near the bottom of the list, demonstrating the utility of adding some mechanisms from CMR to improve model fits. Standard CMR is at the bottom of the list, demonstrating that some of CRU's mechanisms are critical to fitting serial recall data. When CRU's recall termination rule is incorporated into CMR, this improves the model's performance, but it still falls far short of the best CRU model. The most successful model is a version of CRU that uses a few CMR mechanisms: adjustable context-to-feature shared support (α) and self-support (δ) associations, and the associative primacy gradient (ϕ_{s}, and ϕ_{d}). This model provides the best fit at the group level as well as for 100% of the individual participants in comparison to baseline CRU and CMR with both CRU's context-based recall termination mechanism and item identification confusability mechanism. Allowing this version of the model to additionally adjust the start-of-list context integration (β_{start}) parameter, to adjust the feature-to-context learning (γ) parameter, or to adjust both parameters, did not improve model fits. The importance of the associative primacy gradient parameters for a good fit is demonstrated by the worse fit of the version of CRU that was not allowed to adjust those parameters. Similarly, the importance of the alpha and delta parameters is demonstrated by the worse fit of the version of CRU that could not adjust those parameters. In the following sections we examine the ability of a subset of these models to capture different benchmark serial recall phenomena, which gives insight into how these model mechanisms give rise to serial recall performance.

Serial Recall Accuracy and Error Rates

In serial recall, participants generally initiate with the first item of the list and rarely exploit recency to guide retrieval. This is seen in the serial recall accuracy curves of Figure 7, which measure how often a participant recalls the correct item in the correct output position (a more stringent measure than the serial position analysis of free recall). Performance starts high and declines steadily across serial positions.

CRU has multiple mechanisms that support this primacy effect in serial recall performance. CRU always initiates recall using the start-of-list context (equivalent to CMR's $\beta_{\text{start}} = 1.0$), and this context is always most associated with the first item in the list. Likewise, CRU's pre-existing encoding contextual drift gradient (β_{dec}) also provide a strong primacy gradient to boost recall accuracy for early-list items. Finally, the letter-identification gradient (higher confusability toward the end of the list) causes better performance in earlier positions than later positions. Correspondingly, CRU fits the serial recall accuracy curves well, as shown in Figure 7 (top row). Enabling CMR's associative primacy gradient (ϕ_{s}, ϕ_{d}) and pre-experimental context-to-feature memory (α, δ) parameters improves the model's overall fit to the data substantially, but this improvement is not clearly seen in the serial recall accuracy analysis Figure 7, where fits only improve slightly.

These outcomes reproduce prior results showing that CMR mechanisms needed to balance primacy and recency in free recall become less important in serial recall when the task enforces a fixed forward order (Logan, 2021; Lohnas, 2024). Because strict serial recall requires reproducing items from the start, it is not necessary to manipulate the β_{start} parameter downward from its already maximal value of 1.

Figure 7

*Serial recall accuracy (SRAC) fits to Logan (2021) serial recall data for list lengths of 5 (**Left Column**), 6 (**Middle Column**), 7 (**Right Column**) of baseline CRU (**Top**), the best performing CRU variant with free pre-experimental context-to-feature memory (α, δ) and CMR-specific primacy gradient (ϕ_s, ϕ_d) parameters (**Middle**), and CMR with its default position-based recall termination mechanism and CRU's item identification confusability mechanism (**Bottom**).*

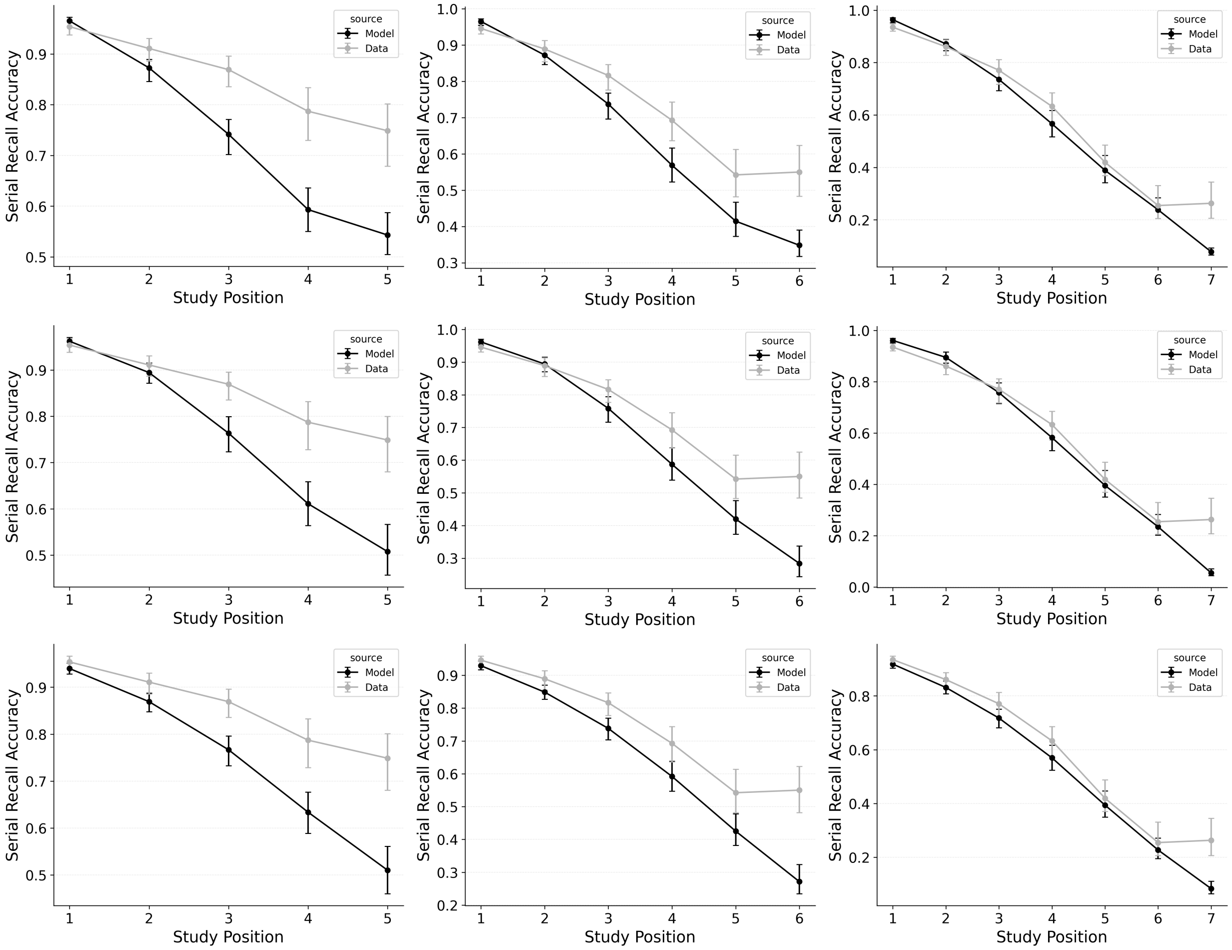
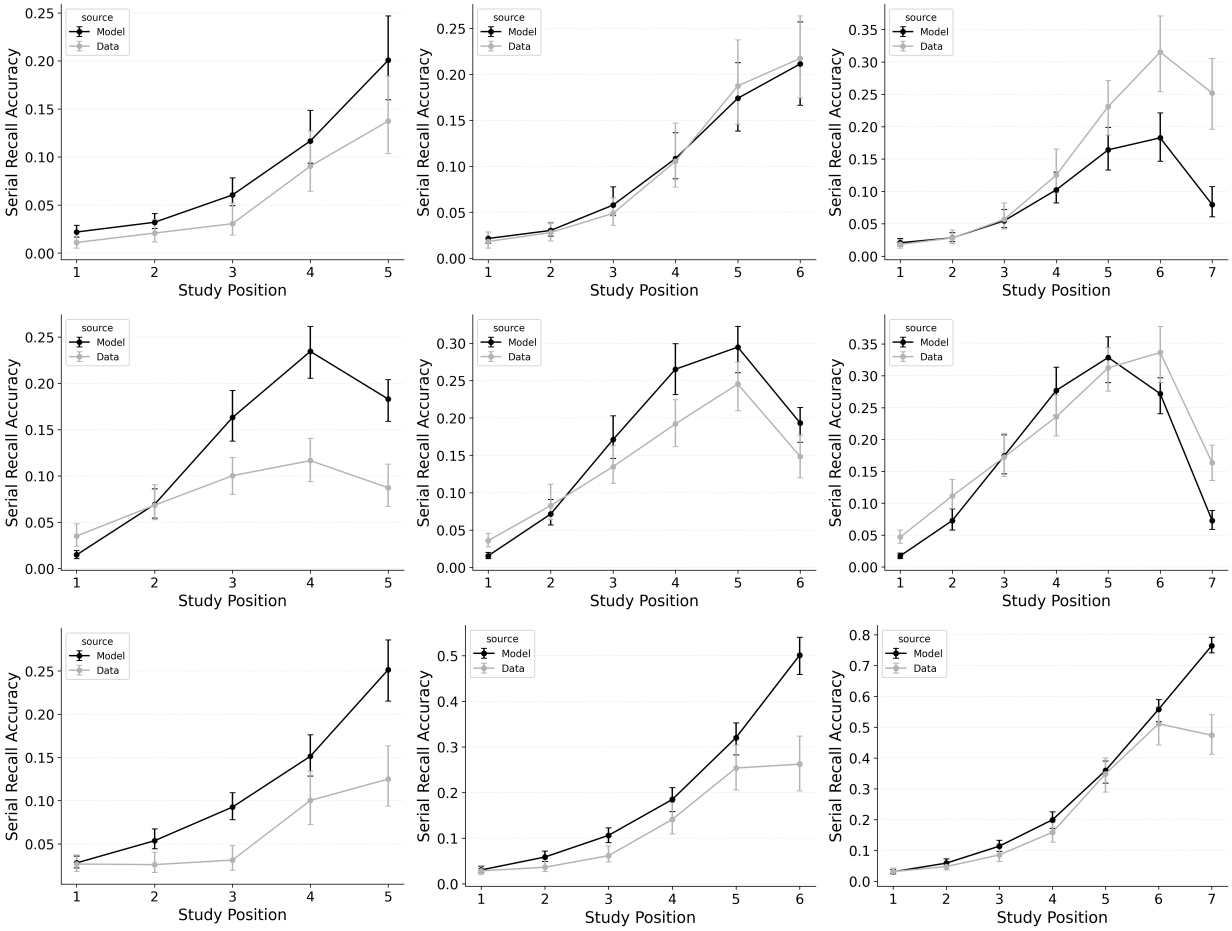


Figure 8

Intrusion, order, and omission error rates (top, middle, and bottom rows respectively) by serial position for list lengths 5, 6, and 7 (left, center, and right columns), in Logan (2021) serial recall data. Lines compare observed error rates with predicted error rates from best performing CRU variant with free pre-experimental context-to-feature memory (α, δ) and CMR-specific primacy gradient (ϕ_s, ϕ_d) parameters.



CRU also explains how different types of response errors are distributed across serial positions. Logan (2021) characterized and simulated three types of errors. Order errors occur when a participant recalls an item from the list but in the wrong position. Intrusion errors occur when a participant recalls an item that was not on the list. Finally, omission errors occur when a participant fails to recall an item that was on the list. Rates of all three types of errors generally increase as a function of serial position in the Logan (2021) dataset. Importantly, intrusion errors are captured by CRU's item identification confusability mechanism. Since CRU can address serial recall accuracy and error patterns without CMR's additional mechanisms, a variant of CMR that drops its pre-existing primacy gradient mechanisms in favor of CRU's confusability mechanism could prove a more parsimonious model of memory search across free and serial recall if the mechanism can be extended to capture confusion rates for words. Figure 8 shows the predicted and observed rates of intrusion, order, and omission errors by serial position for the best performing CRU variant with free pre-experimental context-to-feature memory (α, δ) and CMR-specific primacy gradient (ϕ_{s}, ϕ_{d}) parameters.

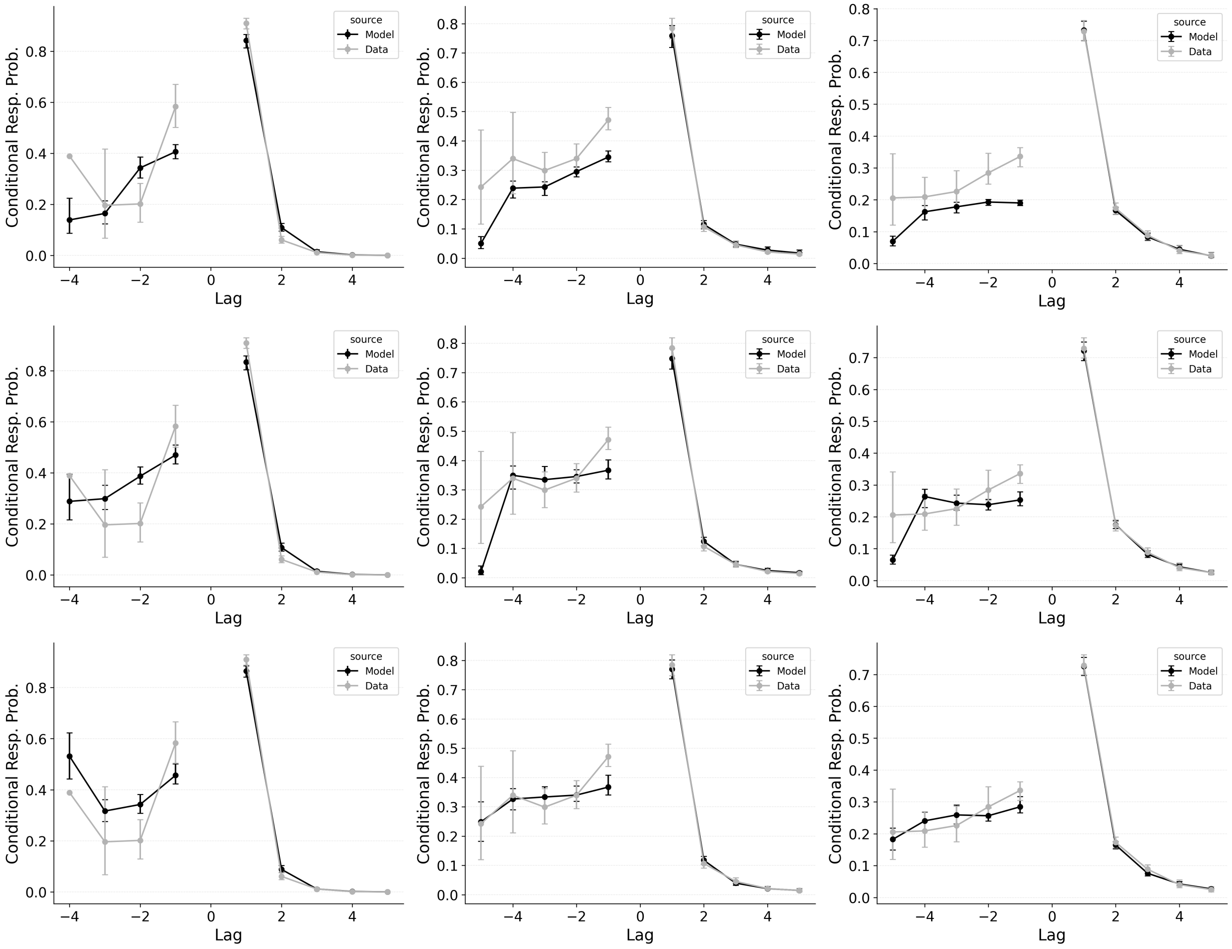
Forward and Backward Transitions in Serial Recall

Serial recall and free recall both show clear temporal structure in participant's sequence of recall transitions, as characterized by the lag-CRP analysis. In strict serial recall, participants typically move forward through the list (i.e., +1 transitions are highly likely), yet small but reliable rates of backward transitions (e.g., fill-in errors) are nonetheless observed. These fill-in errors occur when a participant temporarily skips a position but later "fill in" the item that was missed, effectively jumping backward one or more steps. Just as in free recall, CMR's feature-to-context memory (γ) and pre-experimental context-to-feature associations (α, δ) provide a means for reinstating earlier list context, thereby increasing the model's propensity for backward transitions at short lags. Lohnas (2024) observed this capability in an evaluation of CMR for serial recall, but CRU was found by Logan (2021) to also capture applicable backward transitions in the serial recall task.

Our simulations of the lag-CRP confirm that CMR's mechanisms can help capture these

Figure 9

*Lag-conditional response probability (lag-CRP) fits of baseline CRU (**Top**); best performing CRU variant with free pre-experimental context-to-feature memory (α, δ) and CMR-specific primacy gradient (ϕ_s, ϕ_d) parameters (**Middle**); and CMR with its default position-based recall termination mechanism and CRU's item identification confusability mechanism (**Bottom**) to Logan (2021) serial recall data. Lines compare observed lag-CRP with predicted lag-CRP for the applicable model variant.*



backward transitions Figure 9. Baseline CRU (top row; with $\gamma = 0$ and $\alpha, \delta = 0$) underpredicts the conditional response probability of short backward lags, mirroring its shortfall in free recall. Hence, when CRU is extended to include the same associative mechanisms that CMR uses to promote backward transitions, it better approximates the observed lag-CRP profile in serial recall. Figure 9 (middle row) shows the lag-CRP for the best performing CRU variant with free pre-experimental context-to-feature memory (α, δ) and CMR-specific primacy gradient (ϕ_{s}, ϕ_{d}) parameters.

An Advantage for Context-Based Recall Termination in Serial Recall

In our free recall evaluation, we found that CMR's position-based termination mechanism outperformed CRU's context-based termination mechanism. This is because the latter can lead to premature termination when recall initiates with a strong recency effect, as is common in free recall. However, in our serial recall evaluation, we find that embedding CRU's context-based termination in CMR improves overall model fits, despite dropping two free parameters Table 3. The standard CMR recall termination mechanism underestimates recall rates for later list items, as shown in Figure 7.

In serial recall, recall termination is not a function of the number of recalls made, but rather a function of the current state of the context, as predicted by CRU's unique specification. By tying the end-of-list state to the evolving context, CRU effectively models this behavior with zero additional parameters. In contrast, CMR's exponential growth in stop probability often forces the model either to under- or overshoot the actual stopping point. Resolving this discrepancy may require a more complex model of recall termination that either incorporates insights from both the context-based and position-based mechanisms or that allows for strategic control over recall termination depending on the task at hand.

Discussion

The simulation analyses presented here provide a structured comparison of CRU and CMR in the domain of free recall. CRU is a successful model of serial recall performance (Logan, 2018, 2021; Logan & Cox, 2021). As such, comparing the two frameworks gives insight into how specific model mechanisms contribute to behavioral phenomena that differ between the two tasks. At the outset, we noted how CRU's success in strictly ordered memory tasks might obscure its capacity to handle the broader dynamics of free recall, where retrieval can proceed in many directions and often terminates in flexible ways. By systematically enabling or disabling different mechanisms from CMR excluded from CRU, we showed how features like a dynamic feature-to-context memory, pre-experimental context-to-feature associations, serial position memory strength scaling, flexible recall initiation, and different termination rules can critically shape free recall performance. This analysis reveals that CRU can, in fact, capture many hallmark free recall phenomena when progressively endowed with CMR-like machinery – but in so doing, it gradually converges on CMR's complexity.

By contrast, in our serial recall evaluation, we found that CRU can be substantially improved with the incorporation of certain CMR mechanisms. These mechanisms provide a substantial improvement in overall log-likelihood goodness of fit scores. CRU's streamlined architecture is sufficient to capture key serial recall phenomena, but incorporation of these CMR mechanisms yields small improvements in fit to several phenomena. In contrast to the free recall evaluation, CRU's context-based recall termination mechanism substantially outperformed CMR's position-based termination mechanism. Furthermore, CRU's item identification confusability mechanism was both necessary to capture intrusion errors in serial recall, and sufficient to capture the distribution of order and omission errors across study positions.

Particularly in free recall, the factorial approach underscores the specific changes needed for CRU to handle backward transitions and and robust primacy–recency trade-offs. The strongest free recall benefits benefits emerge when we grant CRU the ability to initiate recall flexibly (β_{start}) using a mix of final and initial study context, a feature that CMR uses to balance primacy and recency effects in recall initiation (Kragel et al., 2015; Morton & Polyn, 2016). CMR additionally implements a proposal by Sederberg et al. (2008) that the primacy effect is supported by increased attention to initial items in study lists. Implementing this associative primacy gradient by allowing the strength of context-to-item associations to scale with serial position (ϕ_{s}, ϕ_{d}) yields appreciable improvements to CRU when paired with β_{start}, suggesting that mechanisms specific to recall initiation and encoding dynamics are crucial for capturing the shape of serial position curve. Broader research focused on patterns of response time distributions in free recall initiation (Osth & Farrell, 2019) suggests that the balance between primacy and recency effects is a key determinant of recall initiation dynamics.

Our results confirm that CMR can be adapted to capture intrusion errors in serial recall by borrowing CRU's confusability mechanism using its g and g_{dec} parameters, simultaneously addressing serial position effects. The mechanism could not evaluated in free recall, as the mechanism depends on empirical perceptual confusability statistics across items, which are not available in word-list free recall. However, the mechanism is likely to be useful in free recall as well, and may be extended to word lists by using a different set of perceptual confusability statistics. Future work should examine the generality of this mechanism across tasks and item sets.

While introducing this flexibility helps CRU capture serial position effects in free recall, additional mechanisms are needed to fully capture the bidirectional lag-contiguity effect, though this is not so much the case in serial recall. Adding pre-experimental associations to CRU's context-to-feature memory (α, δ) provide the flexibility needed to capture forward transitions in the lag-contiguity effect. To also capture the strength of backward transitions, CRU needs to enable feature-to-context learning (γ) so that associations necessary for contacting backward neighbors can be leveraged during retrieval. While addressing backward transitions is crucial for addressing the asymmetric lag-contiguity effect in free recall, such mechanisms may also help address performance in serial recall where probed recall of serial lists (Kahana & Caplan, 2002) and dissociations between forward and backward recall (Li & Lewandowsky, 1993) suggest a similar pattern of associative asymmetry in memory search (Howard & Kahana, 2002).

Prior work examining CMR's ability to capture serial recall phenomena have proposed that free vs. serial recall differences reflect a task-specific configuration of the γ parameter (Lohnas, 2024). Our investigation demonstrates that enabling this mechanism along with pre-experimental context-to-feature memory parameters (α, δ) does improves CRU's ability to simulate backward transitions in serial recall, but also shows that CRU already captures the basic effect without using these mechanisms. Further research is needed to determine whether this dynamic feature-to-context memory is necessary for capturing performance in other memory tasks.

Finally, CRU's method of modeling recall termination was poorly aligned with the variable stopping patterns often observed in free recall. By contrast, CMR's position-based termination mechanism struggled to address recall termination dynamics in serial recall, where participants typically initiate recall with the first item of the list and apparently treat reaching the end of the list as the main cue to terminate retrieval. These results reinforce that a purely context-driven account of termination is better suited to tasks where participants explicitly strive for sequential order. Even in free recall, other research has shown that recall termination probability in depends on factors other than output position. Participants are especially likely to terminate free recall after making an error (Miller et al., 2012; Unsworth et al., 2011), when they are less confident in their responses (Unsworth et al., 2011), when they are less motivated to continue (Dougherty & Harbison, 2007), and when a longer amount of time has passed since the last recall (Dougherty & Harbison, 2007). These results together suggest that recall termination probability is a function of the difficulty of continuing recall, not just the number of recalls made so far. Future work may explain these patterns by relating the accessibility in memory of the next item as predicted by a retrieved-context model to the likelihood of continuing or terminating recall. Moreover, this suggests that the shift from using a leaky competitive accumulation process to a probabilistic termination process was a theoretical step backwards for CMR, as the accumulation process more naturally accounts for the relationship between memory accessibility and successful recall.

By showing that CRU and CMR can be seen as points along a continuum of retrieved-context approaches, our results highlight both the flexibility of retrieved-context theory, and the trade-offs involved in simplifying this theory. CRU's architecture is well-tailored to domains emphasizing forward-ordered retrieval, but omits parameters crucial for capturing the bidirectional and open-ended nature of free recall. Meanwhile, CMR's greater flexibility comes at the cost of added complexity. The results here do not explore potential overfitting, but suggest that each model can be expanded or trimmed depending on task demands. Moreover, CRU's instance-based storage of item–context pairs diverges from CMR's linear associative memory, and the consequences of this difference for model behavior are ambiguous pending further research (J. A. Anderson, 1995; Turner, 2019). While both models have been shown to address item repetitions and out-of-list intrusions (e.g., Logan, 2021; Siegel & Kahana, 2014), it remains unclear which framework better scales across task domains to address the full range of memory phenomena. Comparing serial and free recall performance across a wider range of list lengths and item sets should clarify the practical boundaries of each approach.

The present analysis sidesteps comparison of CRU and CMR's distinct recall competition mechanisms based on the justification that neither model is committed to a specific mechanism for recall competition (Logan, 2021; Morton & Polyn, 2016; Polyn et al., 2009) and simulation using the probabilistic choice rule is more computationally efficient. Nonetheless, with response time distributions providing important constraints for accounts of recall initiation (e.g., Osth & Farrell, 2019) and termination (e.g., Dougherty & Harbison, 2007), the racing diffusion model of recall competition favored in demonstrations of CRU (Logan, 2018, 2021) potentially offers a tractable framework for addressing response time distributions as a theoretical constraint within the likelihood-based fitting approach used here (Tillman et al., 2020). To benefit from this approach, future work should clarify assumptions about when recall competitions begin and end across responses in free and serial recall tasks, and how these assumptions can be tested against response time data (Logan, 2021).

In broader terms, these findings advance our understanding of how context-based models can unify disparate memory phenomena. Indeed, this ambition follows a long line of Gordon Logan's work, which systematically shows that a single cognitive architecture can explain performance across tasks as diverse as visual search, typing, and serial recall. The parallels we draw between CRU and CMR suggest a shared foundation for modeling recall sequences, whether strictly ordered as in serial recall or unconstrained as in free recall. Contemporaneous research indicates that CMR variants can capture core aspects of both free and serial recall (Lohnas, 2024), while CRU's success spans traditional list-learning tasks and broader cognitive-control or motor-sequencing paradigms that rely on forward-chained retrieval (Logan, 2018, 2021). Such breadth underscores the remarkable scope of retrieved-context theory as a unifying explanation across multiple domains, from free recall of word lists to hierarchical production skills. By situating CRU and CMR within a single modeling space, we emphasize how closely related mechanisms can produce seemingly divergent behaviors. This convergence invites researchers to treat free and serial recall findings as complementary constraints on a unified account of memory search, laying groundwork for a more integrated theory of episodic retrieval in both highly directed and more open-ended recall scenarios.

References

Anderson, J. A. (1995). *An introduction to neural networks*. MIT press.

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, *79*(2), 97.

Dougherty, M. R., & Harbison, J. (2007). Motivated to retrieve: How often are you willing to go back to the well when the well is dry? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(6), 1108.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review*, *62*(3), 145.

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). The toronto word pool: Norms for imagery, concreteness, orthographic variables, and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, *14*(4), 375–399.

Healey, M. K., & Kahana, M. J. (2014). Is memory search governed by universal principles or idiosyncratic strategies? *Journal of Experimental Psychology: General*, *143*(2), 575.

Healey, M. K., & Kahana, M. J. (2016). A four-component model of age-related memory change. *Psychological Review*, *123*(1), 23.

Healey, M. K., & Wahlheim, C. N. (2024). PEPPR: A post-encoding pre-production reinstatement model of dual-list free recall. *Memory & Cognition*, *52*(1), 163–181.

Hong, M. K., Gunn, J. B., Fazio, L. K., & Polyn, S. M. (2024). The modulation and elimination of temporal organization in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *50*(7), 1035.

Horwath, E. A., Rouhani, N., DuBrow, S., & Murty, V. P. (2023). Value restructures the organization of free recall. *Cognition*, *231*, 105315.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 923.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299.

Kahana, M. J. (1996). Associative retrieval processes in free recall. *Memory & Cognition*, *24*(1), 103–109.

Kahana, M. J. (2020). Computational models of memory search. *Annual Review of Psychology*, *71*(1), 107–138.

Kahana, M. J., & Caplan, J. B. (2002). Associative asymmetry in probed recall of serial lists. *Memory & Cognition*, *30*(6), 841–849.

Kragel, J. E., Morton, N. W., & Polyn, S. M. (2015). Neural activity in the medial temporal lobe reveals the fidelity of mental time travel. *Journal of Neuroscience*, *35*(7), 2914–2926.

Li, S.-C., & Lewandowsky, S. (1993). Intralist distractors and recall direction: Constraints on models of memory for serial order. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(4), 895.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492.

Logan, G. D. (2018). Automatic control: How experts act without thinking. *Psychological Review*, *125*(4), 453.

Logan, G. D. (2021). Serial order in perception, memory, and action. *Psychological Review*, *128*(1), 1.

Logan, G. D., & Cox, G. E. (2021). Serial memory: Putting chains and position codes in context.

Psychological Review, *128*(6), 1197.

Logan, G. D., & Cox, G. E. (2023). Serial order depends on item-dependent and item-independent contexts. *Psychological Review*.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review*, *108*(2), 393.

Lohnas, L. J. (2024). A retrieved context model of serial recall and free recall. *Computational Brain & Behavior*, 1–35.

Lohnas, L. J., Polyn, S. M., & Kahana, M. J. (2015). Expanding the scope of memory search: Modeling intralist and interlist effects in free recall. *Psychological Review*, *122*(2), 337.

Luce, R. D. (1959). *Individual choice behavior* (Vol. 4). Wiley New York.

Melton, A. W., & Martin, E. (1972). *Coding processes in human memory.*

Miller, J. F., Weidemann, C. T., & Kahana, M. J. (2012). Recall termination in free recall. *Memory & Cognition*, *40*, 540–550.

Morton, N. W., & Polyn, S. M. (2016). A predictive framework for evaluating models of semantic organization in free recall. *Journal of Memory and Language*, *86*, 119–140.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*(5), 482.

Osth, A. F., & Farrell, S. (2019). Using response time distributions and race models to characterize primacy and recency effects in free recall initiation. *Psychological Review*, *126*(4), 578.

Osth, A. F., & Hurlstone, M. J. (2023). *Do item-dependent context representations underlie serial order in cognition? Commentary on logan (2021).*

Polyn, S. M. (2023). Assessing neurocognitive hypotheses in a likelihood-based model of the free-recall task. In *An introduction to model-based cognitive neuroscience* (pp. 303–325). Springer.

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.

Sederberg, P. B., Gershman, S. J., Polyn, S. M., & Norman, K. A. (2011). Human memory reconsolidation can be explained using the temporal context model. *Psychonomic Bulletin & Review*, *18*, 455–468.

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.

Siegel, L. L., & Kahana, M. J. (2014). A retrieved context account of spacing and repetition effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 755.

Storn, R., & Price, K. (1997). Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, *11*(4), 341–359.

Tillman, G., Van Zandt, T., & Logan, G. D. (2020). Sequential sampling models without random between-trial variability: The racing diffusion model of speeded decision making. *Psychonomic Bulletin & Review*, *27*(5), 911–936.

Tulving, E., & Madigan, S. A. (1970). Memory and verbal learning. *Annual Review of Psychology*, *21*, 437–484.

Turner, B. M. (2019). Toward a common representational framework for adaptation. *Psychological Review*, *126*(5), 660.

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2011). Factors that influence search termination decisions in free recall: An examination of response type and confidence. *Acta Psychologica*, *138*(1), 19–29.

Zhang, Q., Griffiths, T. L., & Norman, K. A. (2023). Optimal policies for free recall. *Psychological Review*, *130*(4), 1104.

Table 1

Parameters and structures specifying CRU and CMR.

Symbol	Description	Applicability across models
C	Temporal context layer	Equivalent in both models
F	Item feature layer	Equivalent in both models
M^{FC}	Item-to-context associations	Omitted (or $\gamma = 0$) in CRU
M^{CF}	Context-to-feature associations	Similar in both models
β_{enc} or β_{max}	Encoding context integration rate	Equivalent in both models
β_{dec}	Encoding integration decay	Omitted (or $\beta_{\text{dec}} = 1$) in CMR
β_{rec}	Recall context integration rate	Omitted (or $\beta_{\text{rec}} = \beta_{enc}$) in CRU
β_{start}	Start context integration rate	Omitted (or $\beta_{\text{start}} = 1$) in CRU
α	Shared support	Omitted (or $\alpha = 0$) in CRU
δ	Pre-experimental context-to-item self-support	Omitted (or $\delta = 0$) in CRU
γ	Item-to-context learning rate	Omitted (or $\gamma = 0$) in CRU
ϕ_{s}	Primacy scale	Omitted (or $\phi_{\text{s}} = 1$) in CRU
ϕ_{d}	Primacy decay	Omitted (or $\phi_{\text{d}} = 0$) in CRU
g	Identification sensitivity	Omitted (or $g = \infty$) in CMR
g_{dec}	Identification sensitivity decay	Omitted (or $g = 1$) in CMR
τ	Choice sensitivity	Omitted (or $\tau = 1$) in CRU
θ_{r}	Stop probability growth	Specific to CMR
θ_{s}	Stop probability scale	Specific to CMR

Table 2

Negative log-likelihood (±95% CI) averaged across participants for selected model variants fit to PEERS free recall data. γ*: item-to-context learning rate;* α*: shared support;* δ*: self-support;* ϕ_s*: primacy scale;* ϕ_d*: primacy decay;* β_{start}*: start context integration rate.*

Model Variant	-LL (±95% CI)
CMR (Free γ, α, δ, ϕ_s, ϕ_d, β_start and Position-Based Termination)	587.13 ± 16.84
CRU with Free γ, ϕ_s, ϕ_d, β_start	606.05 ± 16.56
CRU with Free α, δ, ϕ_s, ϕ_d, β_start	608.00 ± 17.00
CMR with CRU’s Context-Based Termination	627.86 ± 17.35
CRU with Free ϕ_s, ϕ_d, β_start	645.04 ± 17.49
CRU with Free β_start	651.35 ± 17.31
CRU	724.01 ± 17.76

Table 3

Negative log-likelihood (±95% CI) averaged across participants for selected model variants fit to the Logan (2021) dataset. γ*: item-to-context learning rate;* α*: shared support;* δ*: self-support;* ϕ_s*: primacy scale;* ϕ_d*: primacy decay;* β_{start}*: start context integration rate. All CRU variants in this table use the context-based end-of-list termination mechanism unless otherwise noted.*

Model Variant	–LL (±95% CI)
CRU with Free α, δ, ϕ_s, ϕ_d	1428.98 ± 222.09
CMR with CRU’s Context-Based Termination	1431.39 ± 223.74
CRU with Free γ, α, δ	1436.94 ± 222.11
CRU with Free ϕ_s, ϕ_d, β_start	1443.14 ± 221.79
CRU	1482.94 ± 230.86
CMR with Own Position-Based Termination Rule	1508.45 ± 208.88