

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Oscillating Echoes: Primary Memory in MINERVA 2

Permalink

<https://escholarship.org/uc/item/5b1172zm>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors

Bull, Adam D

Wong, Roslyn

Reichle, Erik D

Publication Date

2025

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Oscillating Echoes: Primary Memory in MINERVA 2

Adam Bull (adam.bull@mq.edu.au)

School of Psychological Sciences, Macquarie University, NSW 2109, Australia

Roslyn Wong (roslyn.wong@mq.edu.au)

School of Psychological Sciences, Macquarie University, NSW 2109, Australia

Erik D. Reichle (erik.reichle@mq.edu.au)

School of Psychological Sciences, Macquarie University, NSW 2109, Australia

Abstract

Although the MINERVA 2 model provides a detailed account of many findings related to episodic memory, its assumption that primary (working) memory is a single vector of features is unrealistic because it fails to explain the feature-binding problem—how the features of multiple items or objects can be simultaneously maintained but still differentiated. For example, the model cannot explain how the features of “red circle” and “blue square” would be maintained in primary memory in a manner that allows the color features to be correctly associated with their respective shapes. Here we propose a more plausible, biologically inspired implementation of primary memory within MINERVA 2, evaluating its performance using serial-order memory tasks and contextualizing the model’s new assumptions within the broader cognitive science and neuroscience literatures.

Keywords: episodic memory, MINERVA 2, serial-order memory; working memory

Introduction

In this paper, we introduce a new variant of *MINERVA 2* (Hintzman, 1976, 1984, 1986, 1987). This variant addresses a key limitation of the model—that its assumptions about primary or working memory are unrealistic because whatever percepts or concepts are active in primary memory are represented by a single pattern of features. Why is this assumption problematic? Consider a situation in which the model is being used to simulate the paired-associate learning of “red circle” and “blue square,” with the features representing both concepts being coactive in primary memory. In running such a simulation, the standard practice would be to allocate half of the primary memory features to represent one concept and the remaining features to represent the other (e.g., see Hintzman, 1986). This solution, however, begs the question of how the colors remain bound to their respective objects? In other words, there is no sense in which “red” is linked to “circle” or “blue” to “square,” so that the concepts might equally be said to represent “red square” and “blue circle.” The new variant of MINERVA 2 described below addresses this feature-binding problem by drawing inspiration from what is known about the electrophysiology of the brain to implement a quasi-biological version of the model. Before describing this updated version, however, it is first necessary to describe the standard model.

Standard Version of MINERVA 2

MINERVA 2 is a computational model of human memory that simulates judgments of recognition and frequency (Hintzman, 1987, 1988), as well as schema abstraction, associative recall, and categorization (Hintzman, 1986). The model has also been adapted to simulate word identification (e.g., Reichle & Perfetti, 2003), sentence processing (Jamieson & Mewhort, 2009), decision-making heuristics (Dougherty et al., 1999), and as will be reported below, findings related to serial-order memory.

As Figure 1 shows, MINERVA 2 divides memory into two storage systems: *primary memory (PM)* and *secondary memory (SM)*. By this portioning, PM represents the content of one’s conscious, immediate experience and is roughly equivalent to short-term (G. A. Miller, 1956) or working memory (Baddeley & Hitch, 1974), whereas SM is roughly equivalent to long-term memory.

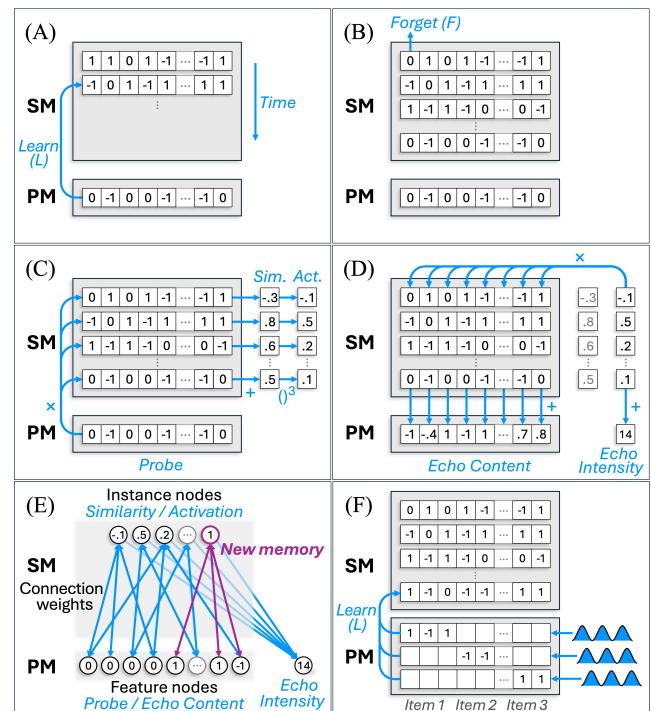


Figure 1. MINERVA 2 model. (A)-(D) Standard model; (E) neural network form; and (F) updated oscillator version.

In the model, experiences are represented by vectors of features that correspond to perceptual and conceptual primitives (e.g., colors, shapes, spatial relations, etc.). Each feature is assigned a value of +1, -1, or 0 to indicate its status (i.e., present, absent, or unknown, respectively).

As Figure 1A shows, during each time step of learning, a vector of features representing the contents of PM is encoded (i.e., copied) into SM as a distinct memory trace. However, encoding is imperfect, with each PM feature having an independent probability, L , of being encoded into SM. And as Figure 1B shows, during each time step, each feature in SM also has an independent probability, F , of being forgotten. These two parameters thus mimic stochastic connectivity in the brain and explain inter-trial variability (Kahana et al., 2018).

As Figure 1C-D show, PM can be used as a probe to retrieve information from SM. This retrieval entails a global matching process in which each feature j of the probe, P_j , is compared in parallel to the features of each memory trace i , T_{ij} , to determine their similarity, S_i , as described by Equation 1. In Equation 1, N_R is the number of active (i.e., non-zero) features in either the probe or trace, such that similarity between the probe and a given trace can range from +1 (identical) to -1 (completely dissimilar).

$$(1) \quad S_i = \sum_{j=1}^N P_j T_{ij} / N_R$$

The activation of each trace i , A_i , is then given by Equation 2. The similarity value of each trace is cubed to increase the signal-to-noise ratio of those traces that are more similar to the probe.

$$(2) \quad A_i = S_i^3$$

Finally, the activated traces are used to generate an “echo” that contains two types of information. The first, *echo intensity*, I , is calculated by summing the activation values across the M traces (see Equation 3). The resulting quantity provides a measure of *familiarity* that has been used to simulate recognition and recency judgments (Hintzman, 1987, 1988).

$$(3) \quad I = \sum_{i=1}^M A_i$$

The second type of information, *echo content*, corresponds to a composite pattern of features where the value of each feature j of the echo content, C_j , is calculated by summing the products of each trace’s activation and its corresponding feature value across the M traces (see Equation 4). The resulting pattern of features is then “cleaned up” by normalizing each feature so that the value of each normalized feature, N_j , ranges between +1 and -1 (see Equation 5). The normalized echo content has been used to simulate associative recall and schema learning (Hintzman, 1986, 1987).

$$(4) \quad C_j = \sum_{i=1}^M A_i T_{ij}$$

$$(5) \quad N_j = C_j / \max(|C_j|)$$

Updated Version of MINERVA 2

Our updated variant of MINERVA 2 adopts the assumption that PM can actively maintain multiple discrete percepts or concepts. By this new account, the contents of PM are represented by separate vectors whose feature values oscillate over time to support their encoding into SM.

By this new account, the activity of the oscillators fluctuates between 0 and 1 using a biologically plausible frequency of 30 Hz (Miller et al., 2018) and a 5-ms phase shift (i.e., slightly out-of-phase). As described by Equation 6, oscillator activity, $y_k(t)$, is calculated every millisecond (t is the index of time) for each representation k in PM using the following parameter values: $A = 0.5$, $B = 30$, $C_k = k + 5$, and $D = 0.5$. Thus, across time, the features associated with each PM representation wax and wane, being concurrently active but separable because each corresponds to a separate pattern of features.

$$(6) \quad y_k(t) = A \sin\left(\frac{2\pi}{1000} Bt - C_k\right) + D$$

During each simulated millisecond, each feature of each PM representation has an independent probability of being encoded, $p(\text{encode})$, into PM, as given by Equation 7.

$$(7) \quad p(\text{encode}) = y_i(t) \times L$$

The final assumption is that the experiencing of any new percept or concept causes a new PM trace to be recruited and “opened” to store the contents of that experience, and the previous memory trace to be “closed.” This opening and closing of SM traces is locked to attention shifts, which are in turn locked to stimulus onset. Therefore, across time, as a new percept or concept is introduced and represented in PM, a new SM trace will be recruited and the features of any percepts or concepts that are currently active in PM will have some probability of being transferred or encoded into SM. And since PM capacity is limited to three representations, the introduction of a new percept or concept will randomly overwrite or replace one of the representations currently being maintained in PM. In the sections that follow, we evaluate the plausibility of these assumptions in a task domain that requires the coordination of primary and secondary memory—serial-order memory.

Serial-order Memory

Serial-order memory refers to the capacity to remember both individual experiences and the order in which they occurred (see Hurlstone et al., 2014). This capacity is one of the first to be studied in experimental psychology (Ebbinghaus, 1885/1913; see Slamecka, 1985) and has often been studied using the *list-recall paradigm*. In this paradigm, participants see or hear a list of items (e.g., words or digits) that are presented sequentially with instructions to recall those items in their correct chronological order (Drewnowski &

Murdock, 1980). Different scoring procedures can then be used to quantify participant performance; for example, *strict serial recall* requires the items to be recalled in their correct temporal order, whereas *lenient serial recall* only requires items to be recalled in their correct relative order (e.g., with omission errors) and *free recall* allows items to be recalled in any order.

Existing accounts of serial-order memory generally require specialized mechanisms to represent the chronological order of items. Examples of these mechanisms include *activation gradients* where relative differences in the activation values of two or more representations can be used to infer their order (e.g., Henson et al., 1996; Page & Norris, 1998), *time tags* that use the output of biological “clocks” (e.g., Burgess & Hitch, 1992) or beginning- and/or end-of-list markers (e.g., Henson, 1998; Logan, 2021) to denote the relative order of two or more representations, and *associative chaining* to “link” representations in their correct chronological order (e.g., Estes, 1991; Lewandowsky & Murdock, 1989). The approach that will be adopted here is a reconstructive one (see Reichle, 1997) in which the contents of the memory traces themselves allow the chronological order of a list of items to be reconstructed during retrieval. How this is done will be described next.

Simulation Method

To evaluate how well our updated version of MINERVA 2 explains serial-order memory, the model was used to simulate one of the classic experiments reported by Drewnowski and Murdock (1980). This experiment required participants to first listen to lists of 3-8 words presented at a fixed rate (667 ms per word) and then recall the words in their correct order. Performance was scored using the three methods discussed earlier: strict serial recall, lenient serial recall, and free recall.

During each simulated trial, each word was assumed to consist of 10 features, with each feature having an equal probability of being +1 or -1. (The features were completely orthogonal across words.) Each list of words also occurred in a unique context of 10 features, with each context feature also having an equal probability of being +1 or -1.

At the start of the trial, the first item and list context would be presented (i.e., loaded into one of the PM vectors supported by an oscillator) for 667 ms. Concurrent with that, a new SM trace would “open” so that any active features in PM would have some independent probability of being encoded into SM (using Equations 6 & 7). The presentation of the next item would then cause it to be loaded into PM (i.e., a PM vector supported by another oscillator), closing the first SM trace and opening a second, thereby allowing any active PM features from either the first or second word to be encoded into SM. This process repeated with each new word presentation, but with the features of the fourth and

subsequent words randomly displacing the features of one of the earlier words.

At recall, the context features were used as the initial probe, with the features of whatever word might be recalled then being used as the next probe, and with the features of any subsequently recalled words being used as subsequent probes. (Recall is described by Equations 1-5, above.) Recall terminated with the first failure to recall a new word.

Because features of multiple words might be generated in the normalized echo content during any given recall attempt, the following scoring procedure was used. First, the Pearson correlation was calculated between the pattern of recalled features (i.e., any features that were not used in the probe) in the normalized echo content and the feature patterns representing each of the words. Second, the word pattern providing the largest correlation was then selected as the word that was recalled. Finally, Drewnowski and Murdock’s (1980) methods were used for scoring strict serial recall, lenient serial recall, and free recall.

The absolute value of any active PM feature that exceeded a “goodness-of-response” threshold, θ , was then set equal to +1/-1, allowing those features to be used as the probe in the next recall attempt. The best-fitting values of this and other model parameters¹ were then used to run simulations with 1,000 simulated participants, as reported below.

Simulation Results

Figure 2 shows the mean proportion of correctly recalled words for (A) strict serial recall; (B) lenient serial recall; and (C) free recall, for lists of 3-8 words. These simulations replicated the classic primacy, recency, and list-length effects, each of which will be discussed in turn.

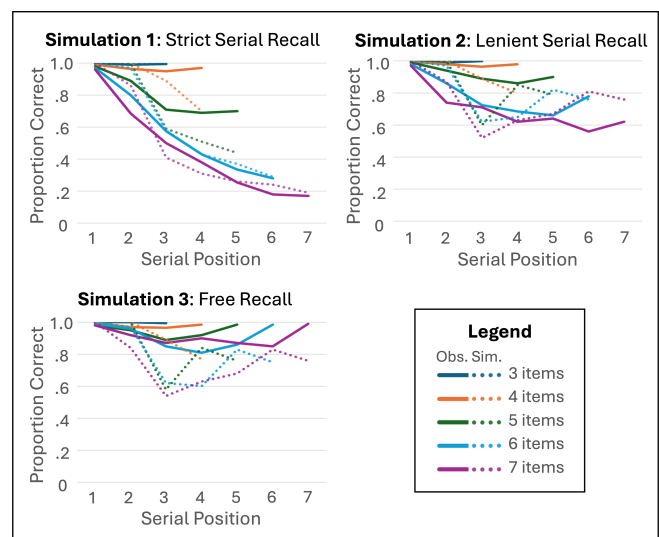


Figure 2: Simulation results.

¹ Two grid searches of the parameter space were conducted to optimize learning rate (L), forgetting rate (F), and “goodness-of-response” threshold (θ) for strict serial recall. Simulated serial position curves were fitted to the observed curves by minimizing

root mean square error (RMSE). The best-fitting parameters from the second search (1,000 simulated subjects per condition) were $L = .17$, $\theta = .64$, and $F = .00$, yielding an RMSE of .0984.

Primacy Effects As Figure 2 shows, the primacy effect, or more accurate recall of words from the beginning of a list, was especially pronounced in strict serial recall, where the decline was monotonic across lists, but was also evident in the other two tasks. Why? Because probing SM with list context tends to retrieve the first word from the list. Although all of the SM traces within a given list share the same contextual features, the normalized echo content generated by the context probe tends to correlate most strongly with the features of the first word because that word is encoded with context in the first SM trace, and because subsequent SM traces are “diluted” by features of subsequent words (which decreases the similarity of the probe to traces of those words by increasing the value of N_R ; see Equation 1). This account of primacy is broadly consistent with the *divided attention* theory of primacy, according to which the addition of items to PM overloads attention and impairs encoding quality in SM (see Azizian & Polich, 2007).

However, our account of primacy is also broadly consistent with an explanation that attributes primacy to the active rehearsal of the initial few words in PM, allowing those words to be better encoded into SM (see Kahana, 2012). Why? Because the first two list words are not at risk of being displaced from PM until its capacity is reached, affording them additional rehearsal time and thus encoding across multiple SM traces. Our model thus also aligns with evidence that faster word presentation rates tend to reduce rehearsal time and primacy (e.g., Glanzer & Cunitz, 1966), as shorter intervals between words allow fewer PM oscillations, decreasing the probability of the words being encoded into SM. Moreover, random item displacement in PM loosely aligns with the random order of rehearsal observed throughout list learning (Rundus, 1971).

List-length Effects Figure 2 indicates that our model replicated the list-length effect, with recall accuracy declining for longer lists. Interestingly, the primacy slope was steeper in longer lists, with the recall of earlier items being impaired by the addition of later items. This interference occurs because the features of early words persist in PM, resulting in additional remote associations with later items and causing those words to be recalled prematurely, out of order. As such, there is more interference towards the middle of a list, and this interference is worse for longer lists.

Because list-length effects reflect how increasing memory load affects the probability of making recall errors, a brief discussion of the nature of these errors is warranted. When recalling ordered lists of words, participants can make four different types of errors: (1) *intrusion errors* in which a word from a previous list is erroneously recalled in the current list; (2) *repetition errors* in which a word is recalled twice; (3) *omission errors* in which a response is not made; and (4) *transposition errors* in which the two words are recalled in their reverse order. Because each trial of our simulations was completed using only a single list of words, intrusion errors were not possible. But the other three types of errors did occur. Repetition errors occurred whenever the normalized

echo content failed to correspond to a new (i.e., not previously recalled) word, which would then terminate the trial, resulting in the omission of any remaining (i.e., unrecalled) words. Transposition errors also occurred due to the fact that, in our model, the associative strength between words in a list tended to decrease with increasing temporal distance; for example, in a 5-word list represented by the letters ABCDE, B and C would be strongly associated due to the fact that their features would be jointly encoded into two or more SM traces, whereas the features of B and E would be only weakly associated due to their features being jointly encoded in a maximum of one SM trace. Such transposition errors would obviously count against performance in strict serial recall but would allow one of the transposed words to be scored as correct in lenient serial recall, and both words to be scored as correct in free recall.

Recency Effects Figure 2 shows that the model exhibited a robust recency effect, or enhanced recall of words near the end of a list, in both lenient serial and free recall. However, contrary to what was reported by Drewnowski and Murdock (1980), this recency effect was not more pronounced for free recall than lenient recall, a distinction often referred to as *immediate recency* and *long-term recency*, respectively (Howard & Kahana, 1999). Evidence suggests that immediate recency occurs in free recall when participants immediately “offload” end-of-list words from primary/working memory (e.g., Glanzer & Cunitz, 1966). Because our model lacks this “offloading” assumption, our simulations are more consistent with *delayed* free recall, where a delay that is introduced between list presentation and recall clears PM and attenuates the size of the recency effect. Any remaining long-term recency benefit is thought to reflect enhanced retrieval from SM (Davelaar et al., 2005). Consistent with this, our model’s decrease in end-of-list errors is likely due to reduced interference from earlier words. Further realism could still be added by incorporating an “offloading” assumption that can reproduce immediate recency, as other models have done previously (e.g., see Kahana, 1996; Reichle, 1997).

Two other points also warrant mention: First, our simulated recency effects also exhibited an unexpected downward inflection at the end of the list, with recall being slightly better for the penultimate than final word. This may be an artifact of the scoring procedure, though the exact cause should be investigated in the future. And second, probing SM with the list context occasionally resulted in the more recent, end-of-list words being recalled first—behavior exhibited by our model and humans in free recall. However, because our free recall simulations did allow a second recall attempt (e.g., using context or a previously recalled word) the model could then attempt to recall words that might have been skipped, as humans can.

Discussion

Biological Plausibility

Although MINERVA 2 explains human memory at the abstract *algorithm-representation* level of analysis (Marr, 1982/2010) and is agnostic about neural *implementation*, it possesses some degree of biological plausibility. For example, there have been several demonstrations of how the model might be implemented as a simple neural network (Reichle et al., 2022; see also Hintzman, 1990; Kelly et al., 2017) using a single set of bidirectional connections between an input/output layer of *feature nodes* (i.e., PM) and a hidden layer of *instance nodes* (i.e., SM; see Figure 1E). During storage, feature nodes that “fire together” in PM “wire together” in SM via an instance node, consistent with Hebb’s (1949) *cell assembly theory*. Each experience is stored separately by recruiting a new instance node from the SM layer and adjusting its unused connection weights according to the activity of the corresponding feature nodes.

Other models (e.g., Nadel et al., 2000; Norman & O’Reilly, 2003) also use this new-node (i.e., multiple-trace) learning assumption to simulate the role of the hippocampus in episodic memory. Instance nodes in MINERVA 2 may conceptually correspond to neurons in the *dentate gyrus* region, which uses sparse representations to store patterns representing individual experiences. By this account, each instance node functions as an index, storing experiences efficiently by linking together features represented elsewhere in the brain (see Teyler & Rudy, 2007).

PM Oscillations

Although oscillatory frequency was set equal to a somewhat arbitrary but realistic value (30 Hz) in our model, it played no functional role because the encoding probability depends only on the area under the curve of the sine waves, which remains constant regardless of frequency. However, evidence suggests that oscillation frequency does play a functional role in PM maintenance (Miller et al., 2018). Electromagnetic field theories even suggest that oscillatory dynamics are central to consciousness and its contents (Hunt et al., 2024).

Likewise, the phase shift plays no functional role in our model. However, evidence suggests that items periodically reactivate out-of-phase and in the same order they were presented in, implying that PM itself maintains serial order (Siegel et al., 2009). This evidence has been exploited by models (e.g., Jensen & Lisman, 2005) which assume a phase-dependent code whereby the representations of items in a list are reactivated sequentially as gamma bursts “riding” larger theta waves. Future iterations could incorporate these serial-order mechanisms into the PM component of our model.

However, it is important to understand that our assumption about the way distinct representations are concurrently maintained in PM by oscillating out-of-phase provides one possible solution to the binding problem (see Figure 1F), and in that capacity, the assumption plays an important role.

PM Indexing

In MINERVA 2, each SM instance node may be conceptualized as corresponding to a hippocampal cell which acts as an index to efficiently combine cortical features into a single memory (see Introduction). Similarly, each PM oscillator node may correspond to a cell or group of cells in the prefrontal cortex which acts as an index to bind cortical features into a single item in experience (Cer & O’Reilly, 2006; Fiebig et al., 2020; Jilk et al., 2008). This mapping makes sense as the prefrontal cortex and hippocampus both play integrative and supervisory roles from “atop” the cortical hierarchy (O’Reilly et al., 2024). The cooperation of these systems, as mapped onto our model, are what allow our model to perform the cognitive operations underlying serial-order memory.

However, it is important to note that neurons in the prefrontal cortex show “mixed selectivity” and participate in multiple, overlapping representations (Miller et al., 2018). Indexing thus might not be as orthogonal as our model suggests, making it useful to explore the consequences of representational overlap in future model implementations.

References

- Azizian, A., & Polich, J. (2007). Evidence for Attentional Gradient in the Serial Position Memory Curve from Event-related Potentials. *Journal of Cognitive Neuroscience*, 19(12), 2071–2081. <https://doi.org/10.1162/jocn.2007.19.12.2071>
- Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- Burgess, N., & Hitch, G. J. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language*, 31(4), 429–460. [https://doi.org/10.1016/0749-596X\(92\)90022-P](https://doi.org/10.1016/0749-596X(92)90022-P)
- Cer, D. M., & O’Reilly, R. C. (2006). Neural mechanisms of binding in the hippocampus and neocortex: Insights from computational models. In H. Zimmer, A. Mecklinger, & U. Lindenberger (Eds.), *Handbook of Binding and Memory: Perspectives from Cognitive Neuroscience* (pp. 193–220). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198529675.003.0008>
- Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: Empirical and computational investigations of recency effects. *Psychological Review*, 112(1), 3–42. <https://doi.org/10.1037/0033-295X.112.1.3>
- Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgments of likelihood. *Psychological Review*, 106(1), 180–209. <https://doi.org/10.1037/0033-295X.106.1.180>
- Drewnowski, A., & Murdock, B. B. (1980). The role of auditory features in memory span for words. *Journal of Experimental Psychology: Human Learning and Memory*, 6(3), 319–332. <https://doi.org/10.1037/0278-7393.6.3.319>

- Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.; pp. viii, 128). Teachers College Press. <https://doi.org/10.1037/10011-000>
- Estes, W. K. (1991). On types of item coding and sources of recall in short-term memory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 155–173). Lawrence Erlbaum Associates, Inc. <https://doi.org/10.4324/9780203772331>
- Fiebig, F., Herman, P., & Lansner, A. (2020). An Indexing Theory for Working Memory Based on Fast Hebbian Plasticity. *eNeuro*, 7(2). <https://doi.org/10.1523/ENEURO.0374-19.2020>
- Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior*, 5(4), 351–360. [https://doi.org/10.1016/S0022-5371\(66\)80044-0](https://doi.org/10.1016/S0022-5371(66)80044-0)
- Hebb, D. O. (1949). *The organization of behavior; a neuropsychological theory* (pp. xix, 335). Wiley.
- Henson, R. N. A. (1998). Short-Term Memory for Serial Order: The Start-End Model. *Cognitive Psychology*, 36(2), 73–137. <https://doi.org/10.1006/cogp.1998.0685>
- Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, 49A(1), 80–115. <https://doi.org/10.1080/027249896392810>
- Hintzman, D. L. (1976). Repetition and Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 10, pp. 47–91). Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60464-8](https://doi.org/10.1016/S0079-7421(08)60464-8)
- Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments & Computers*, 16(2), 96–101. <https://doi.org/10.3758/BF03202365>
- Hintzman, D. L. (1986). “Schema Abstraction” in a Multiple-Trace Memory Model. *Psychological Review*, 93(4), 411–428. <https://doi.org/10.1037/0033-295X.93.4.411>
- Hintzman, D. L. (1987). Recognition and recall in MINERVA 2: Analysis of the “recognition-failure” paradigm. In *Modelling cognition* (pp. 215–229). John Wiley & Sons.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95(4), 528–551. <https://doi.org/10.1037/0033-295X.95.4.528>
- Hintzman, D. L. (1990). Human Learning and Memory: Connections and Dissociations. *Annual Review of Psychology*, 41(1), 109–139. <https://doi.org/10.1146/annurev.ps.41.020190.000545>
- 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), 923–941. <https://doi.org/10.1037/0278-7393.25.4.923>
- Hunt, T., Jones, M., McFadden, J., Delorme, A., Hales, C. G., Ericson, M., & Schooler, J. (2024). Editorial: Electromagnetic field theories of consciousness: opportunities and obstacles. *Frontiers in Human Neuroscience*, 17. <https://doi.org/10.3389/fnhum.2023.1342634>
- Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin*, 140(2), 339–373. <https://doi.org/10.1037/a0034221>
- Jamieson, R. K., & Mewhort, D. J. K. (2009). Applying an exemplar model to the artificial-grammar task: Inferring grammaticality from similarity. *The Quarterly Journal of Experimental Psychology*, 62(3), 550–575. <https://doi.org/10.1080/17470210802055749>
- Jensen, O., & Lisman, J. E. (2005). Hippocampal sequence-encoding driven by a cortical multi-item working memory buffer. *Trends in Neurosciences*, 28(2), 67–72. <https://doi.org/10.1016/j.tins.2004.12.001>
- Jilk, D. J., Lebiere, C., O'Reilly, R. C., & Anderson, J. R. (2008). SAL: An explicitly pluralistic cognitive architecture. *Journal of Experimental & Theoretical Artificial Intelligence*, 20(3), 197–218. <https://doi.org/10.1080/09528130802319128>
- Kahana, M. J. (1996). Associate retrieval processes in free recall. *Memory & Cognition*, 24(1), 103–109. <https://doi.org/10.3758/BF03197276>
- Kahana, M. J. (2012). *Foundations of human memory* (pp. xiii, 349). Oxford University Press.
- Kahana, M. J., Aggarwal, E. V., & Phan, T. D. (2018). The variability puzzle in human memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(12), 1857–1863. <https://doi.org/10.1037/xlm0000553>
- Kelly, M. A., Mewhort, D. J. K., & West, R. L. (2017). The memory tesseract: Mathematical equivalence between composite and separate storage memory models. *Journal of Mathematical Psychology*, 77, 142–155. <https://doi.org/10.1016/j.jmp.2016.10.006>
- Koppelaar, L., & Glanzer, M. (1990). An examination of the continuous distractor task and the “long-term recency effect.” *Memory & Cognition*, 18(2), 183–195. <https://doi.org/10.3758/BF03197094>
- Kwantes, P. J., & Mewhort, D. J. (1999). Modeling lexical decision and word naming as a retrieval process. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale*, 53(4), 306–315. <https://doi.org/10.1037/h0087318>
- Lewandowsky, S., & Murdock, B. B. (1989). Memory for Serial Order. *Psychological Review*, 96(1), 25–57. <https://doi.org/10.1037/0033-295X.96.1.25>
- Logan, G. D. (2021). Serial order in perception, memory, and action. *Psychological Review*, 128(1), 1–44. <https://doi.org/10.1037/rev0000253>
- Marr, D. (2010). *Vision: A Computational Investigation into the Human Representation and Processing of Visual*

- Information. The MIT Press.
<https://doi.org/10.7551/mitpress/9780262514620.001.0001>
- Miller, E. K., Lundqvist, M., & Bastos, A. M. (2018). Working Memory 2.0. *Neuron*, 100(2), 463–475.
<https://doi.org/10.1016/j.neuron.2018.09.023>
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81–97.
<https://doi.org/10.1037/h0043158>
- Nadel, L., Samsonovich, A., L. R., & M. M. (2000). Multiple trace theory of human memory: Computational, neuroimaging, and neuropsychological results. *Hippocampus*, 10(4), 352–368.
[https://doi.org/10.1002/1098-1063\(2000\)10:4<352::AID-HIPO2>3.0.CO;2-D](https://doi.org/10.1002/1098-1063(2000)10:4<352::AID-HIPO2>3.0.CO;2-D)
- Norman, K., & O'Reilly, R. (2003). Modeling hippocampal and neocortical contributions to recognition memory: A complementary-learning-systems approach. *Psychological Review*, 110(4), 611–646. <https://doi.org/10.1037/0033-295X.110.4.611>
- O'Reilly, R. C., Munakata, Y., Frank, M. J., & Hazy, T. E. (2024). *Computational Cognitive Neuroscience* (5th ed.). Wiki Book. <https://compcogneuro.org/>
- Page, M., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, 105(4), 761–781. <https://doi.org/10.1037/0033-295X.105.4.761-781>
- Reichle, E. D. (1997). *A reconstructive theory of serial-order memory* [Doctoral dissertation, University of Massachusetts Amherst]. <https://www.proquest.com/docview/304357449/abstract/7C57EC4D0E5E4AE1PQ/1>
- Reichle, E. D., & Perfetti, C. A. (2003). Morphology in Word Identification: A Word-Experience Model That Accounts for Morpheme Frequency Effects. *Scientific Studies of Reading*, 7(3), 219–237.
https://doi.org/10.1207/S1532799XSSR0703_2
- Reichle, E. D., Veldre, A., Yu, L., & Andrews, S. (2022). A neural implementation of MINERVA 2. In J. Culbertson, A. Perfors, H. Rabagliati, & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society* (Vol. 44, pp. 2278–2284). Cognitive Science Society.
<https://escholarship.org/uc/item/9wd7c291>
- Rundus, D. (1971). Analysis of Rehearsal Processes in Free Recall. *Journal of Experimental Psychology*, 89(1), 63–77.
<https://doi.org/10.1037/h0031185>
- Siegel, M., Warden, M. R., & Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proceedings of the National Academy of Sciences*, 106(50), 21341–21346.
<https://doi.org/10.1073/pnas.0908193106>
- Slamecka, N. J. (1985). Ebbinghaus: Some associations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(3), 414–435.
<https://doi.org/10.1037/0278-7393.11.3.414>
- Teyler, T. J., & Rudy, J. W. (2007). The hippocampal indexing theory and episodic memory: Updating the index. *Hippocampus*, 17(12), 1158–1169.
<https://doi.org/10.1002/hipo.20350>