Memory Retrieval from First Principles

M. Katkov,[1] S. Romani,[2] and M. Tsodyks[1,3,*]
[1]Department of Neurobiology, Weizmann Institute of Science, Rehovot 76000, Israel
[2]Howard Hughes Medical Institute, Janelia Research Campus, Ashburn, VA 20147, USA
[3]Center for Theoretical Neuroscience, Columbia University, New York, NY 10027, USA
*Correspondence: misha@weizmann.ac.il
http://dx.doi.org/10.1016/j.neuron.2017.03.048

The dilemma that neurotheorists face is that (1) detailed biophysical models that can be constrained by direct measurements, while being of great importance, offer no immediate insights into cognitive processes in the brain, and (2) high-level abstract cognitive models, on the other hand, while relevant for understanding behavior, are largely detached from neuronal processes and typically have many free, experimentally unconstrained parameters that have to be tuned to a particular data set and, hence, cannot be readily generalized to other experimental paradigms. In this contribution, we propose a set of "first principles" for neurally inspired cognitive modeling of memory retrieval that has no biologically unconstrained parameters and can be analyzed mathematically both at neuronal and cognitive levels. We apply this framework to the classical cognitive paradigm of free recall. We show that the resulting model accounts well for puzzling behavioral data on human participants and makes predictions that could potentially be tested with neurophysiological recording techniques.

Despite thousands of years of astronomical observations, a clear picture unifying the motion of celestial objects appeared only after the formulation of a very few basic principles, known as Newton's laws. Those principles seemingly appeared from nowhere, but had a great unifying power. Most importantly, they provided the intellectual framework in which different observable phenomena were reduced to the underlying interaction of idealized objects via a small number of different forces. The exact nature of the forces is not specified in Newton's laws, allowing for the application of the laws to a great many different physical phenomena. Progress in physics, as in other scientific disciplines, is achieved by formulation of first principles, development of a mathematical framework for their description, and experimentally testing resulting predictions. It is still unclear whether similar principles exist in neuroscience, in particular those allowing a link between basic components — neurons, synapses — and cognitive functions, like memory, emotions, or language. Here, we propose a set of first principles for memory retrieval and confront them with a classical paradigm in cognitive psychology for studying episodic memory: free recall.

First principles include the following:

(1) The encoding principle states that an item is encoded ("represented") in the brain by a specific group of neurons in a dedicated memory network. When an item is retrieved ("recalled"), either spontaneously or when triggered by an external cue, this specific group of neurons is activated.
(2) The associativity principle states that, in the absence of sensory cues, a retrieved item plays the role of an internal cue that triggers the retrieval of the next item.

The first principle reflects the distributed nature of item encoding, whereas the second one provides a crucial link between neuronal and cognitive processes. As in the case of Newton's laws, the detailed nature of representations and an exact rule for retrieval transitions are not precisely specified, with different instantiations leading to various observable phenomena. In particular, the representations mentioned in the first principle could be random or have more complex structures reflecting the organization of items used in experiments. The precise rules that define the associativity of the second principle could reflect the similarity of neuronal representations for pairs of items and the strength of connections between them. In this contribution, we show how one can apply the proposed set of principles to a concrete cognitive process for which classical experimental results are available.

Free Recall Paradigm

In the standard free recall paradigm, human participants are presented with lists of randomly assembled words and are then asked to recall as many words as possible in an arbitrary order, either immediately after the presentation or after a certain delay. Surprisingly, the task is quite challenging for most people, such that already for very short lists of five or even fewer words, they begin to omit some of the words (Binet and Henri, 1894; Murdock, 1960, 1962; Roberts, 1972; Standing, 1973; Murray et al., 1976). In other, less studied paradigms, participants are not presented with any material but are asked to recall it based on some criteria, e.g., words that begin with a particular letter (Murray, 1975) or objects of a certain class (Graesser and Mandler, 1978). In order to illustrate our basic principles, we first focus on the aspects of the task that are common to all experimental paradigms, namely the issue of recall performance: the number of items that can be recalled (N_r) when all items to be recalled constitute a list of length L. Several reports in the literature presented intriguing results on a power-law relation between the average N_r and L with the exponent of approximately one-half. Interestingly, similar scaling was reported both for the classical free recall paradigm of randomly selected lists of words (Murray et al., 1976) and when participants recall words that

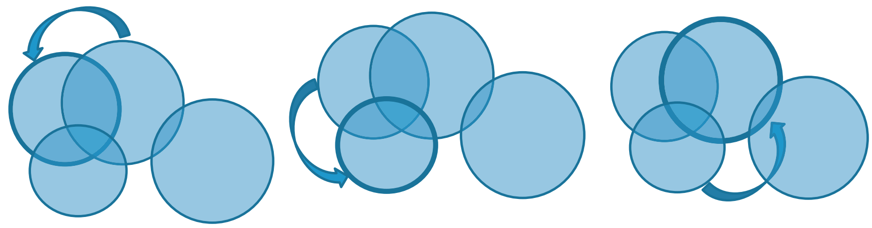

Figure 1. Three Steps of Retrieval Process
Each memory item in a list of four is represented by a randomly chosen population of neurons, here illustrated by blue circles. The size of the circle reflects the number of neurons representing a given item; the size of the overlap between the circles (darker area) reflects the number of neurons representing two given items. An arrow shows one retrieval step from one recalled item to the next chosen according to the size of the overlaps, as explained in the text.

begin with different letters, from A to Z (Murray, 1975). The fact that similar recall performance was obtained for two very different paradigms suggests common computational principles underlying recall, but no unified account of these results was presented despite a large literature on free recall modeling.

Established Models of Free Recall

Trying to model the free recall paradigm illustrates well the dilemma described in the Summary. The phenomenon of recall is a high-level cognitive task that presumably involves many contributing processes, including encoding of memory items in long-term memory, their association to the list to be recalled, and finally the recall from memory itself. None of these processes is currently well understood enough to allow constructing reliable computational models, let alone quantitatively precise ones. It thus appears that a bottom-up design of a biologically realistic model of free recall that would match the experimental results would be a hopeless task. Indeed, the only attempt of realistic neural network modeling of free recall that we are aware of fails to even qualitatively reproduce the correct behavior of N_r with L (Hasselmo and Wyble, 1997). On the other extreme, there are several psychology-driven phenomenological models of free recall (Raaijmakers and Shiffrin, 1980; Howard and Kahana, 2002; Laming, 2009; Polyn et al., 2009; Lehman and Malmberg, 2013). These models involve multiple processes hypothesized to underlie the recall and are specifically designed for the standard paradigm with prior presentation of material in the form of a list, with little relevance for other paradigms. For the standard paradigm, several interesting regularities in recall order of presented lists were reported, in particular recency (tendency to recall better words in the end of the presented list, especially for immediate recall, Murdock, 1962), primacy (same for words at the beginning of the list, Murdock, 1962), and contiguity (tendency to recall consecutively proximally positioned words from the presented list, Kahana, 1996).

The influential model called "search of associative memory," or SAM by Raaijmakers and Shiffrin (1980), postulates several processing stages such as long-term stores, associative retrieval cues, short-term stores, and sampling stage that together determine which words will be recalled and at what time. The resulting model was shown to produce a good fit to the data, including the number of words recalled and the cumulative recall timing, as well as recency and primacy effects, for lists of 10–40 words. Another highly influential model called "temporal context model" (TCM) by Howard and Kahana (2002) postulates the presence of a separate network that stores time-dependent context representations of the list that are evoked by just-recalled words and then serve as retrieval cues for the next word. The main goal of the TCM is to reproduce the temporal contiguity of free recall reported by Kahana (1996). Since the models have quite a complicated structure, they are characterized by a relatively large number of free parameters that cannot be directly related to independent measurements and, thus, have to be tuned to account for the data. As a result, while the fit to the data is excellent (Raaijmakers and Shiffrin, 1980; Polyn et al., 2009), parameters are tuned to the specific range of L s, and increasing the range will require additional tuning. It is thus hard to predict whether the model is compatible with the power-law relation between N_r and L described above. Moreover, it is unclear how these models could be implemented in neural networks. Attempts in this direction lead to a complicated network structure, which does not offer an immediate understanding of retrieval mechanisms (Grossberg and Pearson, 2008).

A New Model

As discussed above, the bottom-up approach to such high-level cognitive tasks as free recall is unrealistic due to insufficient knowledge of its neuronal underpinnings, and psychologically motivated algorithmic models are rather complicated, are detached from neuronal implementation, and don't generalize well for data that were not used for parameter tuning. Here, we present a new, much simpler model based on the two principles suggested above, which are concretized in a way that can be directly tested with electrophysiological recordings (Romani et al., 2013): (1) for encoding, each memory item is represented in the memory network of size N by random neuronal population, such that each neuron in the network is assigned to a given memory with small probability f, independent of other neurons; (2) for associativity, recall is a recursive process of retrieval that is driven by the overlaps between representations of different items, i.e., a set of neurons that belong to representations of two given items (see Figure 1). We emphasize that the overlaps referred to in (2) are between long-term representations, i.e., they are inherent property of memory items to be recalled, rather than acquired during presentation. The first assumption is clearly a simplification because it ignores the complex organization of knowledge in memory but seems reasonable when items that are to be recalled lack an apparent structure, such as, e.g., randomly assembled lists of words or words that begin with a certain letter.

We further assume that the first item to be recalled is chosen at random; subsequently, each next item to be recalled is chosen to be the one from within the list that has the largest overlap with the current item recalled, excluding the item that was recalled in the previous step.

Figure 1 illustrates how the proposed algorithm works for a list of four items, three of which are recalled, after which the retrieval

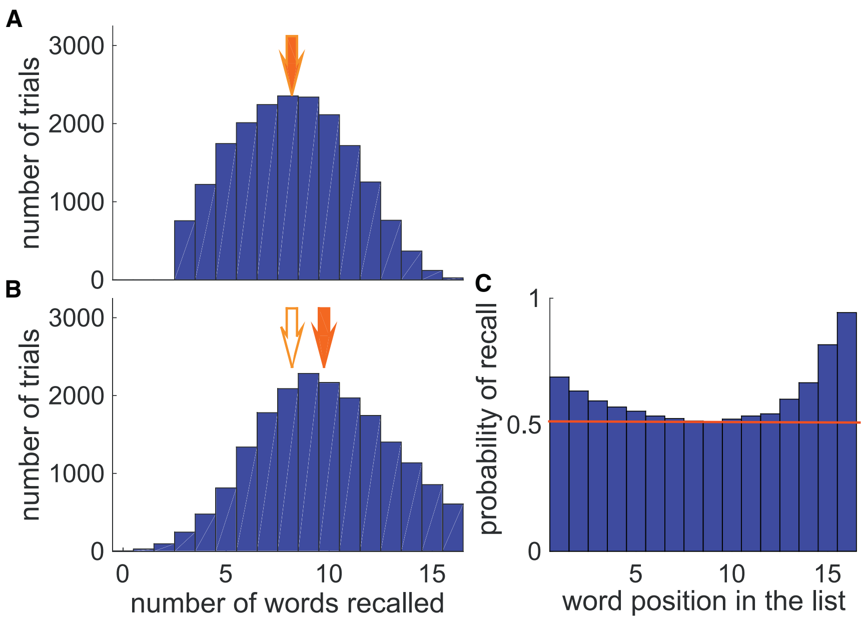

Figure 2. Distribution of Recall Performance: Model versus Data

(A) Distribution of N_r (number of words recalled) obtained by simulating the model 20,000 times. Filled red arrow points to the average N_r that equals 8 in agreement with analytical calculations of the model for $L = 16$.

(B) Same distribution computed for experimental data collected in the lab of Prof. Michael Kahana at the University of Pennsylvania (Miller et al., 2012). Filled red arrow, average number of words recalled equal 9.5; empty red arrow, average number of words recalled after eliminating recency and primacy effects (see text).

(C) Probability of recall for words with different position in the list (serial position curve) obtained from the data. Horizontal red line, serial position curve in the model.

cycles over the same items, leaving the fourth one unretrieved. The process thus described is deterministic apart from the choice of the first item that is taken to be completely random. We point out at the outset that the proposed scheme does not account for any of the regularities that are specific for the classical free recall paradigm described above (primacy, recency, and contiguity). With this limitation in mind, we emphasize that the algorithm is based on a reasonable, common assumption about memory representations and is very simple, characterized by only one parameter, namely the average fraction of neurons in the network representing a given item (sparseness). Mathematically, the first principle is implemented by memory representations as random binary vectors of length N: $\{\xi_i^{\mu} = 0; 1\}$, where $i = 1, \ldots, N$ is an index of a neuron in the memory network, and $\mu = 1, \ldots, L$ is an index of a memory item in the list to be recalled. Each component of each vector is independently chosen to be one with probability f, which is the parameter that defines the sparseness of the representation. The second principle (associativity of the retrieval) is realized with the similarity matrix between different representations defined as $S^{\mu\nu} = \sum_{i=1}^{N} \xi_i^{\mu} \xi_i^{\nu}$, i.e., the number of neurons that encode both items μ and ν. Electrophysiological and computational evidence (Miyashita, 1988; Tsodyks and Feigel'man, 1988; Quiroga et al., 2005; Leutgeb et al., 2007; Gelbard-Sagiv et al., 2008; Barak et al., 2013) indicate that the sparseness parameter f is small, not exceeding a few percentage points.

Theoretical Results and Comparison with Data

Despite the simplicity of the assumptions, the model exhibits several features that were observed in experimental data and makes further predictions.

Scaling Law

Figure 1 illustrates how the proposed retrieval process results in some memories being left out of the recall. The exact order and the number of items recalled is fully determined by the pattern of overlaps between the long-term representations and the choice of the first item to be recalled. Since we assume that the representations are random, the number of items recalled (N_r) for a given list length (L) is also a random quantity. As we showed in Romani et al. (2013), the average performance can be analytically computed to be

$$<N_r> \sim L^{\alpha}; \quad \alpha = \frac{1}{2}\frac{1-f}{1+f}.$$

This result shows that the model accounts for the power-law scaling of performance observed experimentally, with an exponent that depends on the sparseness of long-term memory representations, which is the only parameter of the model. Moreover, the exponent approaches one-half in the limit of sparse coding ($f \to 0$), in agreement with Murray (1975) and Murray et al. (1976) (see Romani et al., 2013 for more detailed comparison with free recall data). The pre-factor in the scaling relationship can also be estimated analytically in the sparse coding limit (see Romani et al., 2013), resulting in the following extremely simple result:

$$<N_r> \approx 2\sqrt{L}.$$

Note that in this limit, the model does not have any parameters left! In other words, based on very basic assumptions about the long-term memory encodings in neuronal networks and the algorithm for item retrieval, the model provides a "first-principled" and parameter-free analytical estimate for recall performance for an arbitrary list length. By numerical simulations of multiple trials, one can go beyond the average performance and estimate the distribution of performances. Here, we compare the model behavior with analysis of the large dataset obtained in the lab of Professor Michael Kahana at the University of Pennsylvania (Miller et al., 2012). In this experiment, participants were engaged with multiple immediate free recall trials with randomly assembled lists of 16 words, with each list presented only once. Simulating the model with $L = 16$ for the same number of trials as in the data, results in the N_r distribution, shown in Figure 2A,

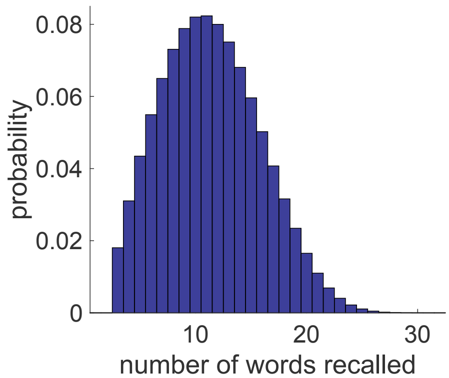

Figure 3. Predicted Distribution of Number of Words Recalled for Lists with Length $L = 32$

contrasted with the distribution obtained from the dataset of Kahana and colleagues (Figure 2B). The distributions have broadly similar shapes, but the experimental average performance is higher than predicted from the model (9.5 versus 8). One could accept this discrepancy as tolerable given the extreme simplicity of the model. A closer look, however, reveals that most of the difference is accounted for by the recency effect observed in the data. Indeed, since the participants started the recall immediately after the exposure to the list, the last few words in the list presumably remained in working memory and, hence, had a significantly higher probability to be recalled (see Figure 2C for the serial position curve, measuring the probability to recall a word versus its serial position in the presented list). Similarly, the first few words in the list also have higher probability of recall, possibly due to rehearsal. In our model, the presentation order has no bearing on retrieval; hence, the predicted serial position curve is flat at the level of 0.5 (horizontal red line, Figure 2C). Recency gradually disappears with increasing delays between acquisition and recall, and primacy is disrupted by preventing rehearsal (Howard and Kahana, 1999). By examining the serial position curve, we therefore conjecture that the probability to recall a word would then converge to the plateau in the curve, approximately one-half; hence, the average number of words recalled would be close to eight, as the model predicts. Eliminating recency and primacy would result in a performance distribution quite similar to that of the model. We conclude that excluding the temporal regularities of the recall that are specific to the free recall paradigm and the trials where chunking or chaining strategies were applied, the model predicts the average performance and its distribution for a list of $L = 16$ surprisingly well.

Inherent Difficulty of Recall

Another theoretical result concerns the inherent ease or difficulty to recall different words. Simulating multiple recalls of randomly assembled lists of items from a large pool, we found that each item has a specific probability to be recalled that depends on the number of neurons representing this item in long-term memory (Romani et al., 2013; Katkov et al., 2015). One consequence of this result is that each word has an inherent "recall probability" that should be largely conserved across participants, if we assume that the distribution of representation sizes over different words is similar for different people. In Katkov et al. (2015), we found that, indeed, the distribution of recall probabilities over words is strongly correlated between different groups of participants.

Order of Recalled Items

The model makes a prediction related to the typical order in which the words are retrieved in a free recall experiment. When a word is recalled, for instance, the first word among the recalled ones, the size of the overlaps between the neuronal representation of this word and the others determines the next word to be retrieved, as described above. There are, on average, larger overlaps with words that have larger representations, which will make them more likely to be retrieved next. The prediction is then that an easy-to-recall word will be recalled earlier compared to a difficult word. This prediction has been confirmed in Katkov et al. (2015).

Correlation between Number of Recalled Words and Their Aggregated Difficulty

An additional, somewhat counterintuitive, prediction of the model is that retrieval of easy words leads to a reduction of the overall number of retrieved words. The reason behind this is, once again, related to the size of neural encodings. Words with larger representations are more likely to be retrieved and more likely to cue the retrieval of other words with large representations; the retrieval process would then quickly enter into a short cycle among the easy items, thereby suppressing the retrieval of small (difficult) words. We have confirmed the existence of a negative correlation between the average recall probability of words in the recalled list and the number of recalled items (Katkov et al., 2015).

Discussion and Further Predictions

We proposed a retrieval algorithm that is based on two simple principles: that memory items are represented by random sparse sets of neurons in dedicated memory networks and that items are retrieved associatively one by one. The first assumption is a common one in theoretical neuroscience; the second one is specific to our model but is common in cognitive literature (Rumelhart and McClelland, 1986; Kropff and Treves, 2007; Russo et al., 2008; Akrami et al., 2012; Lerner et al., 2012; Hinton and Anderson, 2014; Rolls and Deco, 2015). The most surprising result of our analysis is that, at least for one list length ($L = 16$), the model predicts the statistics of recall performance rather well, even though there is not a single parameter to tune. One obvious way to test the model with behavioral methods is to perform experiments with other list lengths. Here, we present the model simulations for $L = 32$ as an example (Figure 3). Another surprising prediction of the model is that each word has an intrinsic recall probability—some words are easier to recall compared to others. Moreover, recalling easy words statistically suppresses the recall of difficult words.

As for the second of Newton's laws, where different forces lead to different observable phenomena, different assumptions about the representation of the items and details of associative processes will lead to different aspects of retrieval. In particular, more realistic encoding and more biologically plausible models for associative retrieval will lead to the description of other cognitive phenomena. For example, in Katkov et al. (2014), we report

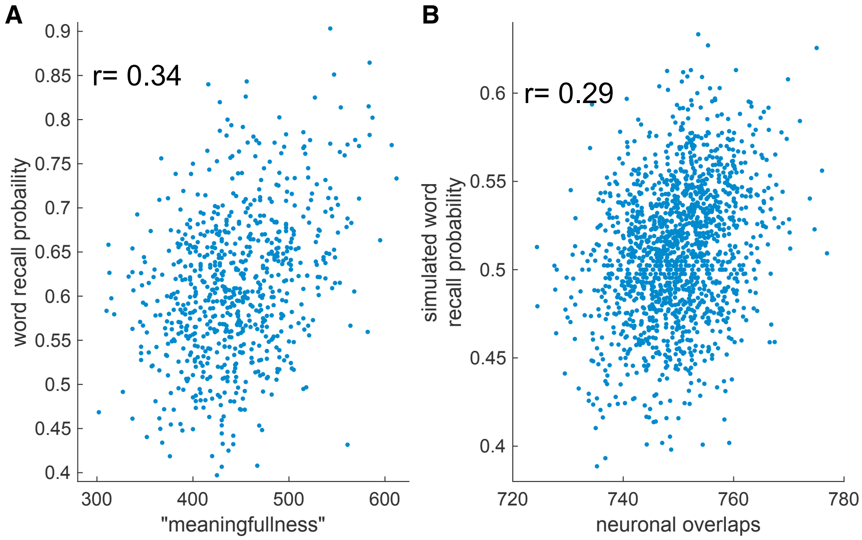

Figure 4. "Meaningfulness" and Recall Probability
(A) Experimentally observed correlation between word recall probability and psychologically defined "meaningfulness" measure, r = 0.34.
(B) Correlations between item recall probability and the total number of neurons for a word that participate in encoding other words. If the neuron encodes three words, it is counted twice.

correlation between intrinsic word recall probability and word length and reproduce this observation by assuming that the variance of words' representation size is related to their syllabic length. Here, we report that the recall probability is also correlated with another word measure introduced in psychological literature, called "meaningfulness," which characterizes how many intrinsic associations with other words a given word has (see Figure 4A; Toglia and Battig, 1978; Coltheart, 1981; Wilson, 1988). Inspired by our model assumptions, we hypothesize that the neuronal underpinning of meaningfulness is the total size of overlaps for a given word with other words in the pool. Analyzing the simulation results, we then found that indeed meaningfulness is correlated with recall probability, similarly to experimental observations (Figure 4B). To isolate the effect of the overlaps, we forced the representation sizes to be the same for each item, which is a good approximation when coding is very sparse. Further elaboration of the encoding principle could include the hierarchical representation of memory items that would result in the emergence of higher-order units of recall, such as chunks of several words, observed in the same dataset in Romani et al. (2016).

In summary, we suggest that a first-principle approach, together with a relatively simple mathematical formulation, is a viable path toward understanding complex cognitive phenomena. This approach allows for predictions of several novel and nontrivial phenomena. This framework is amenable to extensions toward more biologically realistic mechanisms and representations, with the possibility of an even deeper explanatory power. Moreover, by having principles based on neuronal representations, the predictions can be directly tested by measurements of neuronal activity. For example, direct evidence for the first principle was recently obtained using optogenetic activations of memory engrams (Roy et al., 2016). Further measurements will provide constraints on specific realizations of first principles, shedding light on what one could figuratively call, in the spirit of Newton's laws, "neuronal mental forces."

ACKNOWLEDGMENTS

We are grateful to M. Kahana and his University of Pennsylvania research group for making their raw data publicly available. M.K. and M.T. are supported by the EU-H2020-FET 1564 and Foundation Adelis. S.R. is supported by the Howard Hughes Medical Institute. The code to reproduce all the effects of the models presented in current contribution is available at https://github.com/mkatkov/memoryRetrieval.

REFERENCES

Akrami, A., Russo, E., and Treves, A. (2012). Lateral thinking, from the Hopfield model to cortical dynamics. Brain Res. *1434*, 4–16.

Barak, O., Rigotti, M., and Fusi, S. (2013). The sparseness of mixed selectivity neurons controls the generalization-discrimination trade-off. J. Neurosci. *33*, 3844–3856.

Binet, A., and Henri, V. (1894). Mémoire des mots. L'année psychol *I*, 1–23.

Coltheart, M. (1981). The MRC psycholinguistic database. Q. J. Exp. Psychol. *33A*, 497–505.

Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., and Fried, I. (2008). Internally generated reactivation of single neurons in human hippocampus during free recall. Science *322*, 96–101.

Graesser, I.I.A., and Mandler, G. (1978). Limited processing capacity constrains the storage of unrelated sets of words and retrieval from natural categories. J. Exp. Psychol. Hum. Learn. *4*, 86–100.

Grossberg, S., and Pearson, L.R. (2008). Laminar cortical dynamics of cognitive and motor working memory, sequence learning and performance: toward a unified theory of how the cerebral cortex works. Psychol. Rev. *115*, 677–732.

Hasselmo, M.E., and Wyble, B.P. (1997). Free recall and recognition in a network model of the hippocampus: simulating effects of scopolamine on human memory function. Behav. Brain Res. *89*, 1–34.

Hinton, G.E., and Anderson, J.A. (2014). Parallel Models of Associative Memory: Updated Edition (Psychology Press).

Howard, M.W., and Kahana, M.J. (1999). Contextual variability and serial position effects in free recall. J. Exp. Psychol. Learn. Mem. Cogn. *25*, 923–941.

Howard, M.W., and Kahana, M.J. (2002). A distributed representation of temporal context. J. Math. Psychol. *46*, 269–299.

Kahana, M.J. (1996). Associative retrieval processes in free recall. Mem. Cognit. *24*, 103–109.

Katkov, M., Romani, S., and Tsodyks, M. (2014). Word length effect in free recall of randomly assembled word lists. Front. Comput. Neurosci. *8*, 129.

Katkov, M., Romani, S., and Tsodyks, M. (2015). Effects of long-term representations on free recall of unrelated words. Learn. Mem. *22*, 101–108.

Kropff, E., and Treves, A. (2007). The complexity of latching transitions in large scale cortical networks. Nat. Comput. *6*, 169–185.

Laming, D. (2009). Failure to recall. Psychol. Rev. *116*, 157–186.

Lehman, M., and Malmberg, K.J. (2013). A buffer model of memory encoding and temporal correlations in retrieval. Psychol. Rev. *120*, 155–189.

Lerner, I., Bentin, S., and Shriki, O. (2012). Spreading activation in an attractor network with latching dynamics: automatic semantic priming revisited. Cogn. Sci. *36*, 1339–1382.

Leutgeb, J.K., Leutgeb, S., Moser, M.-B., and Moser, E.I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. Science *315*, 961–966.

Miller, J.F., Kahana, M.J., and Weidemann, C.T. (2012). Recall termination in free recall. Mem. Cognit. *40*, 540–550.

Miyashita, Y. (1988). Neuronal correlate of visual associative long-term memory in the primate temporal cortex. Nature *335*, 817–820.

Murdock, B.B., Jr. (1960). The immediate retention of unrelated words. J. Exp. Psychol. *60*, 222–234.

Murdock, B.B. (1962). The serial position effect of free recall. J. Exp. Psychol. *64*, 482–488.

Murray, D.J. (1975). Graphemically cued retrieval of words from long-term memory. J. Exp. Psychol. Hum. Learn. *1*, 65–70.

Murray, D.J., Pye, C., and Hockley, W.E. (1976). Standing's power function in long-term memory. Psychol. Res. *38*, 319–331.

Polyn, S.M., Norman, K.A., and Kahana, M.J. (2009). A context maintenance and retrieval model of organizational processes in free recall. Psychol. Rev. *116*, 129–156.

Quiroga, R.Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. Nature *435*, 1102–1107.

Raaijmakers, J.G.W., and Shiffrin, R.M. (1980). SAM: a theory of probabilistic search of associative memory. Psychol. Learn. Motiv. *14*, 207–262.

Roberts, W.A. (1972). Free recall of word lists varying in length and rate of presentation: a test of total-time hypotheses. J. Exp. Psychol. *92*, 365–372.

Rolls, E.T., and Deco, G. (2015). Networks for memory, perception, and decision-making, and beyond to how the syntax for language might be implemented in the brain. Brain Res. *1621*, 316–334.

Romani, S., Pinkoviezky, I., Rubin, A., and Tsodyks, M. (2013). Scaling laws of associative memory retrieval. Neural Comput. *25*, 2523–2544.

Romani, S., Katkov, M., and Tsodyks, M. (2016). Practice makes perfect in memory recall. Learn. Mem. *23*, 169–173.

Roy, D.S., Arons, A., Mitchell, T.I., Pignatelli, M., Ryan, T.J., and Tonegawa, S. (2016). Memory retrieval by activating engram cells in mouse models of early Alzheimer's disease. Nature *531*, 508–512.

Rumelhart, D.E., and McClelland, J.L. (1986). Parallel Distributed Processing, Volume 2. Explorations in the Microstructure of Cognition: Psychological and Biological Models (MIT Press).

Russo, E., Namboodiri, V.M., Treves, A., and Kropff, E. (2008). Free association transitions in models of cortical latching dynamics. New J. Phys. *10*, 015008.

Standing, L. (1973). Learning 10,000 pictures. Q. J. Exp. Psychol. *25*, 207–222.

Toglia, M.P., and Battig, W.R. (1978). Handbook of Semantic Word Norms (Erlbaum).

Tsodyks, M., and Feigel'man, M. (1988). The enhanced storage capacity in neural networks with low level of activity. Europhys. Lett. *6*, 101–105.

Wilson, M.D. (1988). MRC psycholinguistic database: machine usable dictionary, version 2.00. Behav. Res. Methods Instrum. Comput. *20*, 6–11.