Running head: UNIFIED MODEL OF EPISODIC MEMORY

A Prospective for a Unified Model of Episodic Memory

Adam F. Osth

The Ohio-State University

Simon Dennis

University of Newcastle

Address correspondence to:

Adam Osth

E-mail: adamosth@gmail.com

Word count: 4805

Abstract

Research in episodic memory has become somewhat fractured in recent years. Many investigations have focused on just one memory task, such as recognition memory, free recall, or serial recall. Despite the previous unification of episodic memory with the global memory models, these models were abandoned due to challenging data that arose in each of the aforementioned tasks. In this chapter, we present a review of each of the key constraints that arose in item recognition, free recall, and serial recall that led to the development of newer memory models. We then outline a prospective for how revisions to the global memory models such as sparse representations, likelihood ratio decisions, long distance associations, and simultaneous output of multiple items could potentially be used to overcome the aforementioned challenges.

A Prospective for a Unified Model of Episodic Memory

"A critical problem of long standing in psychological study of memory is concerned with the relation between recall and recognition. In what sense are they the same, and in what sense are they different?" (Tulving & Watkins, 1973, p. 739). This quote began the well revered Gillund and Shiffrin (1984) article that answered Tulving and Watkins's question in the form of the search of associative memory (SAM) model's unification of free recall and recognition memory. The SAM model, along with the other global memory models such as the theory of distributed associative memory (TODAM: Murdock, 1982), MINERVA 2 (Hintzman, 1988), and the matrix model (Humphreys, Bain, & Pike, 1989), unified episodic memory as a whole. Collectively, the global memory models were extended to a wide variety of domains that were presumed to reflect retrieval from episodic memory, including schema abstraction (Hintzman, 1986), judgments of frequency (Hintzman, 1988), memory for serial order (Lewandowsky & Murdock, 1989), priming (Ratcliff & McKoon, 1988), probability estimation (Dougherty, Gettys, & Ogden, 1999), generation of semantic representations (Kwantes, 2005), lexical decision (Kwantes & Mewhort, 1999), and spoken word recognition (Goldinger, 1998).

Despite the zeitgeist that occurred in the years following the publication of the Gillund and Shiffrin article, researchers have become increasingly focused on particular tasks instead of episodic memory as a whole. Hintzman (2011) criticized this tendency in a recent review of episodic memory research and compared researchers focusing on single tasks to the fable of the blind philosophers feeling different parts of an elephant. Just as observing only the trunk of an elephant will not allow one to understand the elephant, solely focusing on recognition or free recall may paint a misleading characterization of episodic memory.

Hintzman's criticism is somewhat unfair, as a number of current episodic memory models have been applied to multiple tasks. Rich Shiffrin's own REM model (Shiffrin & Steyvers, 1997, 1998) which initially emphasized recognition memory has been applied to a variety of tasks such as cued recall (Diller, Nobel, & Shiffrin, 2001), free recall (Lehman & Malmberg, 2013), perceptual identification (Huber, Shiffrin, Lyle, & Ruys, 2001; Schooler, Shiffrin, & Raaijmakers, 2001), lexical decision (Wagenmakers et al., 2004), and the development of knowledge representations (Nelson & Shiffrin, 2013). Nonetheless, the level of task comprehensiveness of other current episodic memory models is still somewhat limited compared to the previous generation's widespread applicability. With the exception of the grouping model of Farrell (2012), models of free recall and serial recall are applied to one task but not the other despite the similarities between the two tasks, and aside from the REM model of free recall, current recall models have not been applied to recognition memory.

Even when one considers empirical research conducted currently, investigations that directly compare multiple tasks are relatively rare. In previous decades task comparisons were more common. Investigations of effects such as word frequency (Gillund & Shiffrin, 1984; Gregg, 1976), semantic similarity (Kintsch, 1968), presentation modality (Watkins, Watkins, & Crowder, 1974), depth of processing (Craik & Tulving, 1975), spacing of repetitions (Glenberg, 1976), and aging (Craik & McDowd, 1987) commonly involved comparisons across multiple tasks, such as recognition and recall or free and serial recall. The logic of the task comparison is that effects that persist across tasks may be more relevant to episodic memory as a whole. It has become somewhat rare to see such comparisons in recent years despite the continued relevance for theories of episodic memory.

Why has the field become increasingly focused on single memory tasks? While we cannot answer this question with precision, our suspicion is that this change in focus came about due to several findings in recognition memory, free recall, and serial recall that challenged the global memory models and led to their abandonment. Researchers have built models that have overcome these challenges but have often omitted consideration of how these models could be extended to other memory tasks.

In this chapter, we present an overview of the challenges that arose, brief discussions of the current models, along with a prospective for unifying the various memory tasks. Interested readers are encouraged to consult the original publications of these models. In this review, we placed somewhat more emphasis on the serial recall task than the others as memory for serial order has received little attention from long term memory researchers. We hypothesize that these challenges can be accommodated within the framework of the global memory models with only minimal revision to the core assumptions of these models.

Recognition Memory

Recognition memory often appeals to researchers due to the simplicity of the task. Unlike recall tasks which may involve varied and sophisticated retrieval strategies, a typical recognition task only involves judging whether a stimulus was experienced on the study list or not. While not all of the global memory models were extended to the different recall tasks, they were all capable of making recognition judgments.

Possibly the two biggest findings in recognition memory that instigated development of newer recognition memory models were the null list strength effect (Ratcliff, Clark, & Shiffrin, 1990) and the mirror effect (Glanzer & Adams, 1985). Despite the differences between the global memory models, all of the models were challenged by these findings.

What united all of the models is that they shared the conception that interference stems entirely from the content of the studied items in memory. Recognition operated in all of the models by virtue of a *global matching* process in which the cues at test are matched against the contents of memory as a whole, producing a single summed memory strength value that indexes the familiarity of the cues. What was considered a success of these models at the time was the ability to account for the list length effect (Strong, 1912), whereby performance decreases as the length of the study list is increased. Global memory models predict a list length effect because each studied list item exhibits variance in its response to a probe cue as a consequence of spurious overlap between their representations. As more items are studied, the variance of the familiarity distributions for both targets and foils is increased, decreasing discriminability.

An unforeseen consequence of this conception of matching and interference is that the *strength* of the stored studied items causes additional interference. Consider a case in which a subset of studied items might be strengthened by repetition, causing additional storage of these repeated items. This manipulation would have a similar functional effect as extending the length of a study list, as the additional copies of the repeated items add variance to the familiarity distributions, decreasing performance on the non-repeated items. This *list strength* prediction of the global memory models was tested by Ratcliff et al. (1990) and found to be false. In all the experiments they conducted, strenghtening a subset of study list by additional study time or repetition does not impair recognition memory of the other list items. An additional complication is that Ratcliff et al. found list strength effects in both free recall and cued recall.

A thorough investigation of these models by Shiffrin, Ratcliff, and Clark (1990) found that the SAM model could predict a null list strength effect by augmenting it with an additional *differentiation* mechanism. Specifically, this assumption states that as items are strengthened their similarity to the other items decreases, decreasing interference among the list items as strength is increased. While this mechanism was successful in predicting a null list strength effect, there is a simpler solution to this problem.

Part of the motivation behind Shiffrin et al.'s reasoning is that differentiation predicts a dissociation between list strength and list length manipulations. Specifically, differentiation models still predict positive list length effects because interference accumulates as more study list items are added to memory. However, more recent evidence suggests that such a dissociation between effects of list length and list strength might not be present. Dennis and colleagues (Dennis & Humphreys, 2001; Dennis, Lee, & Kinnell, 2008; Kinnell & Dennis, 2011) have noted a number of confounds in list length experiments that have been artifactually contributing to the finding of a list length effect. For instance, when immediate testing is used after both short and long lists, the retention interval is longer for the studied items in the long list, causing reduced performance for the long list relative to the short list for reasons unrelated to interference. When these confounds are controlled, all experiments conducted by Dennis and colleagues have found no effect of list length on discriminability

Dennis and Humphreys (2001) argued that null effects of list length and list strength can be predicted from a model in which the representations of studied items do not overlap with each other. Thus, the variance of the target and foil distributions stays constant as list length or list strength are increased. While Dennis and Humphreys built this assumption into a binomial likelihood ratio model (the bind cue decide model of episodic memory, or BCDMEM), the assumption can be similarly built into any of the aforementioned global memory models if one employs orthogonal or highly sparse representations for the studied items.

The other major challenge to the global memory models in recognition memory was the finding of the mirror effect by Glanzer and Adams (1985). A mirror effect is any manipulation that produces opposite effects on the hit and false alarm rates. Glanzer and Adams identified a number of manipulations that produce this effect, including concreteness, imageability, and meaningfulness, but the most popular is the word frequency effect, in which low frequency words exhibit higher hit rates and lower false alarm rates than high frequency words. While global memory models could capture this effect by assuming that low frequency words are better learned than high frequency words (producing a higher hit rate for low frequency words) in addition to a stricter response criterion for low frequency words than high frequency words (producing a lower false alarm rate), this account is challenged by the fact that the mirror effect persists in two alternative forced choice testing

(Glanzer & Bowles, 1976), which is assumed to not involve a response criterion.

The memory models that succeeded the global memory models, such as REM, BCDMEM, Attention Likelihood Theory (ALT: Glanzer, Adams, Iverson, & Kim, 1993), and the model of McClelland and Chappell (1998), all captured the mirror effect by usage of a likelihood ratio decision mechanism. Rather than simply use the familiarity of a cue as the direct basis for a recognition decision, likelihood ratio models weigh the odds that the cue is a target against the odds that the cue is a foil. Consider a case where a cue elicits low memory strength but comes from a class of items that is expected to have low memorability (a high frequency word, for instance). In this case, it is somewhat ambiguous as to whether the item is a target or a foil. Consider another case where the cue elicits low memory strength but comes from a class that is expected to have high memorability (low frequency words). It is much more likely in this case that the cue is a foil and can be rejected with higher confidence. It is this comparison of memory strength against the expected memorability for the cue that allows the models to produce mirror effects.

While there are many differences among likelihood ratio models, such as the degree to which they're correctly informed about the expected memorabilities of the classes of test items (Criss & McClelland, 2006) and the explanations they employ for word frequency effects, considerable power is derived from the usage of a likelihood ratio transformation in addressing the mirror effect. Recently, Glanzer, Hilford, and Maloney (2009) developed analytic expressions of the likelihood ratio transformation that can be employed using both equal variance and unequal variance familiarity distributions that are Gaussian in shape. Because the familiarity distributions of the global memory models are approximately Gaussian in the limit[1], this transformation can be applied on the back-end of any of the aforementioned global memory models to produce the mirror effect.

Free Recall

While all of the global memory models addressed the recognition memory task, only the SAM model was applied to free recall (Gillund & Shiffrin, 1984; Raaijmakers & Shiffrin, 1981). In a sense, SAM can be viewed as a mechanistic implemention of the Atkinson and Shiffrin (1968) model, as SAM possesses a limited capacity short-term memory buffer that strengthens the associations between its contents in long term memory. That is, if two items A and B are present in the buffer together, the associative strength between A and B in long term memory is a direct function of the time they co-occur in the buffer. Free recall consists of emptying the contents of the buffer if testing immediately follows the study phase and subsequently using the retrieved items as cues for recalling the remaining items.

A major challenge that arose for the SAM model of free recall was the *contiguity effect*. The contiguity effect was discovered by Kahana (1996) in a re-analysis of several free recall datasets. A striking regularity was present in the transition probabilities participants made between words in their recall sequences: after recalling a word from a particular serial position in the study list, participants were likely to recall an item that was studied in close proximity to the just recalled item. The contiguity effect was also asymmetric, in that participants were more likely to recall an item that was studied after the just-recalled item than to recall an item that was studied earlier.

The first regularity is not problematic for the SAM model due to the nature of the short-term memory buffer. Items that are studied in proximity to each other in the study list (such as items from serial positions 5 and 6) are more likely to simultaneously occupy the buffer for longer durations of time than distant items (such as 12 and 15). Thus, if item 5 is recalled, it will have more associative strength to nearby studied items such as the 4th and 6th items than it will to a distant item such as item 15, making proximal transitions more likely than distant ones. The asymmetry in the contiguity effect is more problematic for the basic SAM model as its original instantiation assumed symmetric associations between the list items. That is, as items A and B co-occur in the rehearsal buffer, the associative strength between A and B is equivalent to that between B and A. By modifying this original assumption with asymmetric associations in the forward direction (stronger A-B than B-A associations), Kahana enabled the basic SAM model to predict stronger forward than backward contiguity.

If contiguity effects arise from co-occupation in the short-term store, a key prediction of the SAM model is that contiguity effects should be abolished if intervening distracting activity that is sufficient to empty the contents of the buffer is placed between the list items. This prediction was tested by Howard and Kahana (1999) by placing a demanding arithmetic task between all of the list items for varying degrees of time and examining the resulting contiguity effects. Contrary to the predictions of the SAM model, contiguity effects persisted as the time of the distracter activity increased even up to the duration of 16 seconds. In subsequent investigations, contiguity effects have been found at longer timescales using a final free recall test (Howard, Youker, & Venkatadass, 2008) and a probed recall procedure (Kilic, Criss, & Howard, 2013).

The scale invariance of the contiguity effect led to the development of the temporal context model (TCM: Howard & Kahana, 2002). In TCM, items are associated to a context vector that consists of an exponentially decaying representation of the previous list items. Put another way, TCM consists of asymmetric associations in the forward direction in which the associative strength between the list items is an exponential function of their distance in the list. A critical distinction from SAM's rehearsal buffer is that items never drop out of the context vector; instead their strengths continue to decrease exponentially as items are added to the context vector. It is for this reason that TCM predicts contiguity effects across time scales: items are not expunged from the context vector during intervening distracter activity and thus associations can still be created between items separated in time. TCM also employs orthogonal representations for list items, which we have previously argued are effective in reducing interference among the list items in recognition memory.

While TCM and its variants have been successful in accounting for the effects of recency, contiguity, and semantic similarity (Polyn, Norman, & Kahana, 2009; Sederberg, Howard, & Kahana, 2008), the long distance associations in TCM that enable its success in addressing contiguity effects can be built into the original global memory models. As we will discuss shortly, long distance associations can also support serial recall, although new assumptions about how retrieval operates are required to address data that are challenging for models that employ inter-item associations.

Serial Recall

Serial recall is one of the oldest tasks in the empirical study of memory, dating back to the studies of Nipher (1878) and Ebbinghaus (1885/1913). Despite the historical importance of the task, it has received little attention from long term memory researchers. Of all of the global memory models, only the TODAM model was applied to serial recall (Lewandowsky & Murdock, 1989).

Serial recall operated in TODAM by assuming pairwise associations among the items in a serial list along with start and end markers. That is, for a list of items ABC, the associations start-A, A-B, B-C, and C-end were stored. Recall operated by using the start cue as an attempt to retrieve the first item in the list. The TODAM model is an example of a *chaining* model of serial recall, in that direct associations among the list items are created at encoding, much like creating links among a chain. At retrieval, each retrieved item is used as a cue for its successor to recreate the list in order. The SAM, REM, and TCM models can similarly be considered chaining models of free recall in that they share the assumption that associations among list items are created and retrieved items are used as cues.

A key prediction made by chaining models in serial recall is that an error in recall should result in further errors because the erroneously recalled items are used as cues for further retrievals. In his seminal dissertation, Henson (1996) tested this prediction in a number of experiments and found this prediction to be false. For instance, in a list such as ABCDE, if a subject erroneously skips from A to C during recall, a chaining model with asymmetric associations in the forward direction predicts that recall of D is most likely to follow. Instead, Henson found that in these instances, recall of B is often twice as likely as recall of D. Henson dubbed this the *fill-in effect* because it was as if participants were "filling in" the missing item from their recall sequences. The opposite case, in which recall continues to D from the C item, is called an in-fill error.

A similar effect can also be seen in mixed lists of confusable and non-confusable items. When a list of phonologically confusable items is studied (such as rhyming consonants B, D, P, etc.), serial recall is worse than when a list of phonologically non-confusable items is studied (Conrad & Hull, 1964). A test of chaining models can be constructed by creating an alternating list of phonologically similar and dissimilar items, such as BRDXPY, in which the R, X, and Y letters are non-confusable and the B, D, and P letters are confusable with each other due to their shared phoneme. Chaining models predict that if a participant erroneously begins recall with an item such as P, the subsequent recall is most likely to be an item near the erroneously recalled item P, such as X or Y. Instead, participants are just as likely to recall the correct item R as they would in a list with no confusable items (Baddeley, 1968; Henson, Norris, Page, & Baddeley, 1996). When errors are plotted by output position, they reveal a sawtooth pattern in which errors spike on the confusable items and dip on the non-confusable items.

Henson (1996) argued that the proper model for serial recall is instead a *positional* model, in which items are associated to their positions within the sequence rather than to the other list items. Positive evidence for positional associations comes from Henson's findings that when participants erroneously intrude items from prior lists, they are likely to be recalled in the same output position as their serial position on the studied list (Conrad, 1960). That is, if a participant learns and recalls a list ABCDE and subsequently learns FGHIJ, if the participant were to initiate recall with an item from the previous list, the intrusion is most likely to be the first item from the previous list (A) despite the fact that the last item on the previous list (E) was more recently experienced. Positional models have a natural and intuitive explanation for these data in that different items that occupy the same positions possess similar representations, whereas it is unclear how pure chaining models could produce positional intrusions of this nature. Nearly all current serial recall models consist of item-position associations and lack any direct item-item associations. These include the model of Burgess and Hitch (1999), Henson (1998)'s start-end model, and the oscillator-based associative recall model (OSCAR: Brown, Preece, & Hulme, 2000).

How could global memory models overcome these challenges to chaining models? While one might be inclined to add positional representations to models such as SAM or TODAM to address the serial recall task, this is a rather unappealing solution in light of recent comparisons of free and serial recall. In particular, it has been found that both free and serial recall are similarly affected by word length, articulatory suppression, and presentation rate (Bhatarah, Ward, Smith, & Hayes, 2009), as well as semantic similarity (Golomb, Peelle, Addis, Kahana, & Wingfield, 2008), and the two tasks exhibit similar profiles of contiguity (Bhatarah, Ward, & Tan, 2008). Free recall is identical to serial recall using short lists of items, and even with longer lists, free recall resembles serial recall when recall is initiated with the first item (Ward, Tan, & Grenfell-Essam, 2010). Given the few dissociations between the two tasks, it is parsimonious to assume the two tasks utilize similar representations.

Another problem with positional representations is that there has been no satisfying account as to how the position codes are generated and used. Henson (1998) used position markers that code for position relative to the start and end of the list, but it is not specified as to how participants can anticipate where the end of the list is, especially in cases where the length of the list is unknown to the participant. Another problem concerns how the position codes for each position are reinstated at retrieval to be used as cues; this reinstatement is assumed rather than specified in positional models[2].

There are also findings that are troublesome for positional models and more easily explained with the concept of inter-item associations. When participants study a list of high frequency bigrams, serial recall performance is better than with lower frequency bigrams (Baddeley, 1964). Items presented adjacent to each other on a serial list show better transfer to a paired associate cued recall task than items that are distant from each other (Crowder, 1968), as if participants are transferring the inter-item associations formed from the serial list to the paired associate list. When a serial list contains a repeated element (e.g.: AB<u>C</u>DE<u>C</u>F), after participants recall the first occurrence of the repeated element (C), participants are more likely to recall later elements from the sequence (F) than for control lists with no repeated elements (Wickelgren, 1966). This follows intuitively from chaining models, which predict more errors following a repeated element due to the multiple inter-item associations (both C-D and C-F associations in a list with a repetition).

A simpler approach than using separate item-position and inter-item associations is to approximate positional information from the inter-item associations among the list items. Dennis (2009) demonstrated this using a matrix model of forward asymmetric associations between the list items. Rather than iteratively chaining through the associations among the list items, the entire sequence is retrieved simultaneously and output. The probability of outputting a candidate sequence is determined by comparing the candidate sequence to the stored matrix of associations from the study episode. The more differences there are between a candidate sequence and the stored associations, the less likely the candidate sequence is to be output. It is for this reason that the model predicts a fill-in effect: for a study list ABCDEF, a sequence with a fill-in error such as ACBDEF has only two misses to the original list, whereas a sequence with an in-fill error such as ACDEF has five misses (the B-C, B-D, B-E, and B-F associations). An illustration of a stored ABCDEF matrix along with candidate ACBDEF and ACDEF matrices can be seen in Figure 1. For the purposes of clarity, the figure shows the simple case where associative strength between two list items does not depend on their distance from each other on the list. More realistic assumptions, such as exponentially decaying associative strength over list positions, can also be implemented in the model.

Positional information can be derived from this model simply by assuming that all of the list items have some degree of similarity to each other. Dennis did this by assuming that there was a common component to all of the item vectors (i.e.: a feature that is shared among all the item vectors). Because each item is associated to the list items that preceded it, after learning has taken place each item has different degrees of association to the common component of the item vectors. For a six item list, the first item, which gets associated to the five succeding items, would have five associations to the common component, while the last item, which does not get associated to any of the list items, consequently has no associations to the common component. When prior list items are considered in the candidate sequences at retrieval, they are more similar to the stored associations when they are in the same position as their prior occurrence due to their similar degrees of association to the common component. That is, if list 1 is ABCDE and list 2 is FGHIJ, if C intrudes during output, it is most likely to be in the third output position as it bears the same degree of association to the common component (two associations) as the item that was actually stored in that position (H).

The assumptions of this model could potentially be built into some of the global memory models. For instance, one could store three way associations between context and two items that are asymmetric in the forward direction in a tensor as described by Humphreys et al. (1989). However, as we previously mentioned, another constraint on serial recall is simultaneously capturing free recall performance with the same associative structure. The model employs long distance associations between list items that we previously argued may be critical for capturing free recall performance. While the retrieval of the entire list of items does not resemble typical free recall performance, this method of retrieval is mostly suited for short lists of items. In short lists, free recall bears a strong resemblance to serial recall (strong primacy and weaker recency; Ward et al., 2010). Longer lists of words, in contrast, cannot be easily recalled using the simultaneous list output, as the number of candidate sequences to consider at retrieval increases exponentially with list length. One possibility is that longer lists are more likely to elicit item-by-item chaining in both free and serial recall.

Conclusions

The original global memory models had a strong appeal because they reduced episodic memory to a simple associative framework that could be extended across multiple tasks. Specifically, they described associations among items and context as the basis of long term memory and retrieval success was determined by the match between the cues and the contents of memory. Extensions to tasks other than the typical memory paradigms was straightforward in that one merely had to specify the cues, item representations, and associations in order to derive predictions about performance.

These principles need not be abandoned. Rather, the models may only require minor revisions in order to address the constraints from the major memory tasks that we have described. Specifically, item recognition can be addressed by minimizing interference among the list items and using a likelihood ratio decision mechanism, free recall requires long distance associations among the list items that are stronger in the forward than backward direction, and serial recall requires a retrieval mechanism that approximates positional information from the associations among the list items.

A unified model such as the one we propose would be useful not just in understanding memory retrieval itself, but would have useful psychometric applications as well. For instance, with age performance in item recognition stays roughly constant while performance in associative memory tasks drops considerably (Ratcliff, Thapar, & McKoon, 2011; Old & Naveh-Benjamin, 2008; Craik & McDowd, 1987). Fitting a unified memory model to the data from the different populations might result in interprettable parameter estimates that could inform researchers about the locus of the memory impairment.

One should also note that new empirical work may be required to place each of these memory tasks on common footing. There are a number of methodological differences between the tasks: recall tasks tend to employ short lists of items whereas recognition memory tasks tend to employ relatively long lists. Whereas free recall experiments tends to involve unique stimuli on every trial, serial recall experiments tend to re-use stimuli from trial to trial. An additional complication is that certain regularities in memory tasks may be due to task-specific encoding strategies. In a comparison of serial recall and short term item recognition, Duncan and Murdock (2000) found that serial position functions in either of the tasks changed completely when participants received a memory test opposite of what they were expecting (e.g.: being tested on serial recall when an item recognition test was expected). Empirical data comparing multiple tasks may enable researchers to better understand how encoding differences are relevant do task differences and can also provide enormous constraint on a unified model.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (pp. 89–195). New York: Academic Press.

Baddeley, A. (1964). Immediate memory and the "perception" of letter sequences. *Quarterly Journal of Experimental Psychology*, *16*, 364-367.

Baddeley, A. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology*, *20*, 249-264.

Bhatarah, P., Ward, G., Smith, J., & Hayes, L. (2009). Examining the relationship between free recall and immediate serial recall: Similar patterns of rehearsal and similar effects of word length, presentation rate, and articulatory suppression. *Memory & Cognition*, *37*(5), 689–713.

Bhatarah, P., Ward, G., & Tan, L. (2008). Examining the relationship between free recall and immediate serial recall: The serial nature of recall and the effect of test expectancy. *Memory & Cognition*, *36*(1), 20–34.

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, *107*, 127–181.

Burgess, N., & Hitch, G. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*, 551-581.

Conrad, R. (1960). Serial order intrusions in immediate memory. *British Journal of Psychology*, *51*, 45–48.

Conrad, R., & Hull, A. (1964). Information, acoustic confusion and memory span. *British Journal of Psychology*, *55*.

Craik, F. I. M., & McDowd, J. M. (1987). Age differences in recall and recognition.

Journal of Experimental Psychology: Learning, Memory, and Cognition, *13*(3), 474–479.

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*(3), 268–294.

Criss, A. H., & McClelland, J. L. (2006). Differentiating the differentiation models: A comparison of the retrieving effectively from memory model (REM) and the subjective likelihood model (SLiM). *Journal of Memory and Language*, *55*, 447–460.

Crowder, R. G. (1968). Evidence for the chaining hypothesis of serial verbal learning. *Journal of Experimental Psychology*, *76*(4), 497–500.

Dennis, S. (2009). Can a chaining model account for serial recall? In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the XXXI Annual Conference of the Cognitive Science Society* (pp. 2813–2818).

Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review*, *108*(2), 452–478.

Dennis, S., Lee, M. D., & Kinnell, A. (2008). Bayesian Analysis of Recognition Memory: The Case of the List-Length Effect. *Journal of Memory and Language*, *59*, 361–376.

Diller, D. E., Nobel, P. A., & Shiffrin, R. M. (2001). An ARC-REM model for accuracy and response time in recognition and recall. *Journal of Experimental Psychology-Learning Memory and Cognition*, *27*(2), 414–435.

Dougherty, M. R. P., Gettys, C. F., & Ogden, E. E. (1999). MINERVA-DM: A memory processes model for judgements of likelihood. *Psychological Review*, *106*(1), 180–209.

Duncan, M., & Murdock, B. B. (2000). Recognition and recall with precuing and postcuing. *Journal of Memory and Language*, *42*, 301–313.

Ebbinghaus, H. (1885/1913). *Memory.* New York: Dover. (H.A. Ruger & C. E. Bussenius Trans. Original work published 1885)

Farrell, S. (2012). Temporal clustering and sequencing in short-term memory and episodic memory. *Psychological Review*, *119*(2), 223–271.

Gillund, G., & Shiffrin, R. M. (1984). A Retrieval Model for Both Recognition and Recall. *Psychological Review*, *91*(1), 1–67.

Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory and Cognition*, *13*(1), 8–20.

Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review*, *100*(3), 546–567.

Glanzer, M., & Bowles, N. (1976). Analysis of the word-frequency effect in recognition memory. *Journal of Experimental Psychology: Human leanring and memory*, *2*(1), 21–31.

Glanzer, M., Hilford, A., & Maloney, L. T. (2009). Likelihood ratio decisions in memory: Three implied regularities. *Psychonomic Bulletin & Review*, *16*(3), 431–455.

Glenberg, A. M. (1976). Monotonic and nonmonotonic lag effects in paired-associate and recognition memory paradigms. *Journal of Verbal Learning and Verbal Behavior*, *15*, 1–16.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251–279.

Golomb, J. D., Peelle, J. E., Addis, K. M., Kahana, M. J., & Wingfield, A. (2008). Effects of Adult aging on utilization of temporal and semantic associations during free and serial recall. *Memory & Cognition*, *36*(5), 947–956.

Gregg, V. H. (1976). Word frequency, recognition, and recall. In J. Brown (Ed.), *Recall and recognition.* London: Wiley.

Henson, R. N. A. (1996). *Short-term memory for serial order.* (Unpublished doctoral dissertation. MRC Applied Psychology Unit, University of Cambridge, Cambridge, England.)

Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology*, *36*, 73–137.

Henson, R. N. A., Norris, D., Page, M., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology*, *49A*, 80–115.

Hintzman, D. L. (1986). Schema Abstraction in a Multiple-Trace Memory Model. *Psychological Review*, *93*(4), 411–428.

Hintzman, D. L. (1988). Judgments of Frequency and Recognition Memory in a Multiple-Trace Memory Model. *Psychological Review*, *95*(4), 528–551.

Hintzman, D. L. (2011). Research strategy in the study of memory: Fads, fallacies, and the search for the "coordinates of truth". *Perspectives on Psychological Science*, *6*(3), 253–271.

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.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*, 268–299.

Howard, M. W., Youker, T. E., & Venkatadass, V. (2008). The persistence of memory: Contiguity effects across several minutes. *Psychonomic Bulletin & Review*, *15*(1), 58–63.

Huber, D. E., Shiffrin, R. M., Lyle, K. B., & Ruys, K. I. (2001). Perception and preference in short-term word priming. *Psychological Review*, *108*(1), 149–182.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different Ways to Cue a Coherent Memory System - a Theory for Episodic, Semantic, and Procedural Tasks. *Psychological Review*, *96*(2), 208–233.

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

Kilic, A., Criss, A. H., & Howard, M. W. (2013). A causal contiguity effect that persists across time scales. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(1), 297–303.

Kinnell, A., & Dennis, S. (2011). The list length effect in recognition memory: An analysis of potential confounds. *Memory & Cognition*, *39*, 348–363.

Kintsch, W. (1968). Recognition and free recall of organized lists. *Journal of Experimental Psychology*, *78*(3), 481–487.

Kwantes, P. J. (2005). Using context to build semantics. *Psychonomic Bulletin and Review*, *12*, 703–710.

Kwantes, P. J., & Mewhort, D. J. K. (1999). Modeling lexical decision and word naming as a retrieval process. *Canadian Journal of Experimental Psychology*, *53*, 306–315.

Lehman, M., & Malmberg, K. J. (2013). A buffer model of memory encoding and temporal correlations in retrieval. *Psychological Review*, *120*(1), 155–189.

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review*, *96*, 25–57.

McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review*, *105*(4), 724–760.

Murdock, B. B. (1982). A Theory for the Storage and Retrieval of Item and Associative Information. *Psychological Review*, *89*(6), 609–626.

Nelson, A. B., & Shiffrin, R. M. (2013). The co-evolution of knowledge and event memory. *Psychological Review*, *120*(2), 356–394.

Nipher, F. E. (1878). On the distribution of errors in numbers written from memory. *Transactions of the Academy of Sciences of St. Louis*, *3*, CCX–CCXI.

Old, S. R., & Naveh-Benjamin, M. (2008). Differential Effects of aging on item and associative measures of memory: A meta-analysis. *Psychology and Aging*, *23*, 104–118.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2009). A Context Maintenance and Retrieval Model of Organizational Processes in Free Recall. *Psychological Review*, *116*(1), 129–156.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of Associative Memory. *Psychological Review*, *88*(2), 93–134.

Ratcliff, R., Clark, S. E., & Shiffrin, R. M. (1990). List-Strength Effect : I. Data and Discussion. *Journal of Experimental Psychology: Learning Memory and Cognition*, *16*(2), 163–178.

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review*, *95*(3), 385–408.

Ratcliff, R., Thapar, A., & McKoon, G. (2011). Effects of Aging and IQ on Item and Associative Memory. *Journal of Experimental Psychology: General*, *140*(3), 464–487.

Schooler, L., Shiffrin, R. M., & Raaijmakers, J. G. W. (2001). A model for implicit effects in perceptual identification. *Psychological Review*, *108*, 257–272.

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

Shiffrin, R. M., Ratcliff, R., & Clark, S. E. (1990). List-Strength Effect: II. Theoretical Mechanisms. *Journal of Experimental Psychology: Learning Memory and Cognition*, *16*(2), 179–195.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM - retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*(2), 145–166.

Shiffrin, R. M., & Steyvers, M. (1998). The effectiveness of retrieval from memory. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 73–95). London: Oxford University Press.

Strong, E. K. J. (1912). The effect of length of series upon recognition memory. *Psychological Review*, *19*, 447–462.

Tulving, E., & Watkins, M. J. (1973). Continuity between recall and recognition. *American Journal of Psychology*, *86*, 739–748.

Wagenmakers, E. J., Steyvers, M., Raaijmakers, J. G. W., Shiffrin, R. M., van Rijn, H., & Zeelenberg, R. (2004). A model for evidence accumulation in the lexical decision task. *Cognitive Psychology*, *48*(3), 332–367.

Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the relationship between free recall and immediate serial reacll: The effects of list length and output order. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1207–1241.

Watkins, M. J., Watkins, O. C., & Crowder, R. G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning and Verbal Behavior*, *13*, 430–447.

Wickelgren, W. A. (1966). Associative intrusions in short-term recall. *Journal of Experimental Psychology*, *72*, 853–858.

Notes

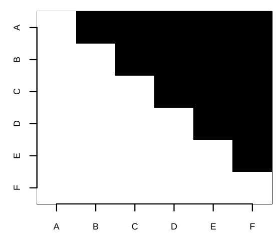
[1]Because the global memory models derive familiarity by summing over the memory strengths of a large number of memories, this sum will be well characterized by a Gaussian distribution by virtue of the central limit theorem.

[2]A partial solution to the reinstatement problem came from the OSCAR model (Brown et al., 2000), in which position codes come from a bank of temporal oscillators which synchronize to the items on the list. Recall is initiated by resetting the temporal oscillators to the state that began the list. As the oscillators rotate through the positions they held during the study phase they are used to cue the list items. While an advantage of the oscillator framework is that it does not require reinstatement of the positions for every list item, it still remains unclear as to how the oscillators could be reset to their initial position.

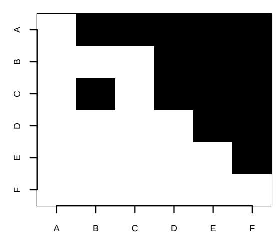
Figure Captions

Figure 1. Matrices representing an ABCDEF list, an ACBDEF list (fill-in error), and an ACDEF list (in-fill error). The ACBDEF list is more similar to the ABCDEF list than the ACDEF matrix.

ABCDEF



ACBDEF



ACDEF