Running head: PUTTING STM INTO CONTEXT

Putting short-term memory into context: Reply to Usher and colleagues (2008)

Michael J. Kahana

Department of Psychology

University of Pennsylvania

Philadelphia, PA 19104, USA

Per B. Sederberg

Department of Psychology and

Center for the Study of Brain, Mind, and Behavior

Princeton University

Princeton, NJ 08540, USA

Marc W. Howard

Department of Psychology

Syracuse University

Syracuse, NY 13244, USA

Abstract

The temporal-context model posits that search through episodic memory is driven by associations between the multiattribute representations of items and context. Context, in turn, is a recency weighted sum of previous experiences or memories. Because recently processed items are most similar to the current representation of context, Usher and colleagues suggest that the temporal-context model (*TCM-A*) embodies a distinction between short-term and long-term memory, and that this distinction is central to *TCM-A*'s success in accounting for the pattern of recency and contiguity observed across short and long time scales. We dispute the claim that context in *TCM-A* has much in common with classic interpretations of short-term memory. The idea that multiple representations interact in the process of memory encoding and retrieval (across time scales), as proposed in *TCM-A*, is very different from the classic dual-process view that distinct rules govern retrieval of recent and remote memories.

Putting short-term memory into context: Reply to Usher and colleagues (2008)

In their comment on our article, Usher, Davelaar, Haarmann, and Goshen-Gottstein (in press) argue that *TCM-A*'s successful account of short- and long-term recency, contiguity, and their dissociations, depends on its implicit use of a short-term memory store that is distinct from long-term memory. Their central argument is that a distinction between short-term and long-term memory is necessary to explain dissociations between immediate and long-term recency.

Before turning to the specific critiques of *TCM-A* raised by Usher and colleagues, it is important to note that they do not dispute the central role of a time-varying context signal for explaining long-term recency and long-range contiguity effects in free recall (Howard & Kahana, 1999, 2002). Furthermore, they do not dispute *TCM-A*'s ability to account for many of the key dissociations between immediate and long-term recency using a single retrieval mechanism. Rather, they argue that these successes arise from *TCM-A*'s employment of separate short-term and long-term memory mechanisms. Although the distinction between weight-based associative memory and an activation-based context representation has always been a central feature of *TCM*, we dispute the claim that this distinction (as embodied in *TCM-A*) is synonymous with the classic division of short-term and long-term memory (e.g., Atkinson & Shiffrin, 1968; Raaijmakers & Shiffrin, 1980, 1981; Sirotin, Kimball, & Kahana, 2005).

Usher and colleagues present several specific criticisms of *TCM-A*. Most notably, they argue that whereas people are able to discriminate items on the current list from items on earlier lists, the context mechanism in *TCM-A* fails to support such list discrimination. They also argue that *TCM-A* fails to account for several specific properties of recency functions including end-of-list distractor effects, list length effects, the shape of the recency effect itself, and interactions between recency and presentation rate. We first discuss each of these critiques of *TCM-A* and then conclude with a discussion of the relation between *TCM-A* and dual-process models of free recall.

List discrimination

Without presenting any simulations, Usher and colleagues claim that *TCM-A* cannot distinguish items on one list from items of previous lists. They alternatively suggest that this "limitation" of the model results from (a) the lack of a list context representation in *TCM-A*, and (b) *TCM-A*'s assumption that contextual drift is driven by item representations. They claim that "*TCM-A* can only be applied to 1-list paradigms, so it cannot truly account for phenomena that require multiple lists..."

Although Sederberg, Howard, and Kahana (in press) focused on single list effects, one can easily show that the contextual dynamics in *TCM-A* can readily distinguish items on the current list from items on prior lists. In *TCM-A*, the representation of context at the time-of-test is a much stronger cue for the current list than it is for earlier lists. This leads the model to predict that recalls will primarily come from the target list, but that participants will occasionally commit prior list intrusions (PLIs). Furthermore, the exponential nature of the contextual drift process predicts that PLIs will largely come from the immediately preceding list, as observed in the data (e.g., Murdock, 1974; Zaromb et al., 2006). *TCM-A* also predicts that PLIs from earlier lists will tend to retrieve contexts from those lists resulting in an increased tendency to commit successive PLIs from the same list, as reported by Zaromb et al. (2006). Table 1 shows simulation results (Lohnas, Polyn, and Kahana, in prep.) obtained using a generalization of *TCM-A* that includes semantic relatedness among items[1]. As can be seen in the table, this model is able to fit the pattern of correct responses, prior list and extralist intrusions, and the distribution of PLIs across trials. Using the same parameters, the model is also able to match the serial position curve, including its “sigmoidal” recency effect (Figure 1a), and the contiguity effect as seen in the conditional-response probability as a function of lag (Figure 1b). A more detailed analysis of the model’s implications for multilist phenomena is currently underway (Lohnas et al., in preparation).

The ability to discriminate items on successive lists is not novel to *TCM-A*’s contextual drift process. As pointed out by Howard and Kahana (2002), with uncorrelated item representations and without retrieved context *TCM-A* closely approximates random context models (e.g., Mensink & Raaijmakers, 1988; Anderson & Bower, 1972; Bower, 1972; Murdock, 1997; Howard & Kahana, 1999). The contextual retrieval process in *TCM-A* will further separate context representations across successive lists, making list discrimination easier rather than more difficult. *TCM-A*’s contextual drift mechanism predicts that list discrimination will be more difficult when lists contain overlapping items because context is composed of item-related information. This prediction has been documented experimentally (e.g., Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Anderson & Bower, 1972).

An interesting open question is whether *TCM-A* can account for people’s ability to recall the list before the last list, as studied by Shiffrin (1970) and more recently by Ward and Tan (2004) and Jang and Huber (2008). Although people have great difficulty with this task, under some conditions their recall of the list before last (list $i-1$) exceeds their intrusion rate from the most recent list (list i). Jang and Huber (2008) found that this ability to selectively target items from list $i-1$ was significantly enhanced when lists i and $i-1$ were separated by a recall period (during which participants recalled items from list $i-2$). In discussing their findings they wrote “As predicted by Howard and Kahana’s (2002) TCM, these results indicate that recall of information from long-term memory serves to drive context change, thus isolating a prior list.”

Although TCM’s contextual drift process during test is indeed consistent with the list-before-last findings, further model development would be needed to capture all aspects of this intriguing paradigm. In particular, one would either need to add a post retrieval recognition process (e.g., Sirotin et al., 2005; Kahana, Dolan, Sauder, & Wingfield, 2005) or a mechanism for temporarily suppressing the context of the most recent list. It remains to be seen whether a list specific context representation is necessary to account for these data.

Elimination of Recency in Delayed Free Recall

Usher et al. (in press) claim that "*TCM-A* predicts an attenuation of recency in delayed free recall, rather than its complete elimination, as a result of its exponential activation-store." They note that experiments employing a sufficiently long distractor period can completely eliminate the recency effect in delayed free recall and that this observation was the "major motivating factor for assuming an abrupt, displacement-based buffer store, in the first place." Usher et al attribute this "failure" of *TCM-A* to the exponential nature of the contextual drift process.

Actually, *TCM-A* can easily eliminate the recency effect in delayed free recall by increasing the drift of the simulated distractor items at the end of the list (Howard, 2004; Howard, Wingfield, & Kahana, 2006). We did not do so in our simulations because our target dataset, which used a 16 second distractor, did not completely eliminate the recency effect in the data. The left panel of Figure 2 shows our simulations of the 30 second delayed free recall condition of Postman and Phillips (1965), which had a distractor long enough to eliminate recency. Clearly, an exponential recency mechanism does not imply recency in delayed free recall.

Usher et al. (in press) further argue that even if the model could predict the lack of recency in delayed free recall, it would fail to predict contiguity when the same distractor is interpolated between list items. The right panel of Figure 2 shows that *TCM-A* predicts a robust contiguity effect in continual-distractor free recall even when simulated with the identical parameters that eliminated recency in the Postman and Phillips (1965) delayed free recall simulation.[2]

Although one may think that the exponential nature of the context drift process directly implies an exponential forgetting process, this is not the case. The form of the recency effect in free recall depends strongly on the dynamics of the recall process. Even though context has an exponentially weighted history of recent items at the end of the list, the act of recalling an item reinstates that item's associated context, which in turn cues the next item, and so forth. In *TCM-A*, the primacy gradient and noise in the accumulators enable the model to occasionally recall pre-recency items early in output. Once an earlier list item is recalled, its retrieved context will tend to cue its neighbors, resulting in a recency function that can be sigmoidal (see Figures 3 and 8 in Sederberg et al., in press).

Tuning of recency and contiguity effects

Usher et al. (in press) point out that in the simulations of continual-distractor free recall reported in Sederberg et al. (in press), recency and contiguity effects appear to be more sharply tuned than in the data. Similarly, in our simulations of amnesia, *TCM-A* showed a less extended recency effect than in the data. These are valid points regarding the specific *TCM-A* implementation presented in Sederberg et al. (in press). Although *TCM-A* can produce more extended recency and contiguity effects under these conditions, we "tied our hands" by trying to fit a large number of empirical facts across multiple experimental conditions with a single set of parameters. This creates a tension in the model: creating more extended recency and contiguity effects would require lowering the variance of the accumulator noise parameter, but this would result in fewer transitions at remote lags because the accumulator noise is the only source of variability in the model.

This is remedied in recent work by Polyn, Norman, and Kahana (submitted) who generalize *TCM-A* to include semantic and source features into the context representation. Variation in the semantic (and source) relatedness among list items produces far more remote transitions and enables the model to simultaneously fit data on temporal, semantic and source clustering, and inter-response times.

List length effects

Usher et al point out that whereas *TCM-A* predicts a dissociation between immediate and long-term recency, the one study that has examined list length effects did not find evidence for such a dissociation (Greene, 1986). This highlights an interesting puzzle insofar as list length manipulations are closely related to proactive interference manipulations (Murdock, 1974). Davelaar et al. (2005) found a dissociation between recency in immediate and continual-distractor free recall for proactive interference, but Greene failed to find such a dissociation for list length.

Although *TCM* does not match the Greene (1986) results, Greene had participants study and free recall pairs of words—a very different methodology from the free recall experiments simulated in Sederberg et al. (in press). Insofar as no one has examined list length effects in continual-distractor free recall of lists of individual items, one cannot know whether *TCM-A*'s predictions are incorrect. A study that simultaneously examines within-list proactive interference effects (i.e., manipulating list length) and across-list proactive interference effects in immediate and continual distractor free recall would enable a far better appraisal of the model.

Sigmoidal Probability of First Recall Functions

In most immediate free recall studies, participants exhibit the highest probability of initiating recall with the last study-item and the probability of first recall (PFR) decreases monotonically with each previous serial position (e.g., Hogan, 1975; Howard & Kahana, 1999; Kahana, Howard, Zaromb, & Wingfield, 2002; Howard, Vankatadass, Norman, & Kahana, 2007; Bhatarah, Ward, & Tan, 2008; Golomb, Peelle, Addis, Kahana, & Wingfield, 2008; Polyn et al., submitted). However, Usher et al. (in press) note two studies that, when reanalyzed by Laming (1999) and Howard and Kahana (1999), were found to exhibit a non-monotonicity whereby the last item was not recalled first with higher probability than the other recency items (Murdock, 1962; Murdock & Okada, 1970). In addition to the two datasets mentioned by Usher et al. (in press), our reanalysis of the data from Murdock and Walker (1969) also reveals a non-monotonic probability of first recall.

Usher et al. (in press) claim that such "sigmoidal" probability of first recall curves support the operation of a short-term buffer because one way to have multiple items equally active at the start of recall is if they are all active in a short-term buffer. Further, they argue that *TCM-A* is unable to produce this pattern because the item activations due to the end-of-list context decrease exponentially into the past.

Given that the vast majority of published studies report monotonic PFR functions, we did not seek to explain the relatively rare observations of non-monotonic effects. Nonetheless, non-monotonic PFR functions could easily arise if rehearsal of items leads to a difference between the nominal serial position and the functional serial position (i.e., the serial position within the sequence of rehearsals, e.g., Brodie and Murdock, 1977) . Consider for example if participants occasionally rehearsed the last 2 or 3 list items after the presentation of the last item and before the recall signal. In this case, *TCM-A* would predict that they would most often initiate recall with the last rehearsed item, which would often differ from the last presented item. Thus, the model could produce non-monotonic PFR functions under conditions that encourage rehearsal. We have no independent evidence that rehearsal differences between the studies can account for the variation in the PFR functions, and indeed, there may be other factors at work. That said, *TCM-A*'s account is consistent with the vast majority of published free recall studies.

Although Usher and colleagues have presented these data as a critical challenge to *TCM-A*, neither Davelaar et al. (2005) nor Usher et al. (in press) have simulated probability of first recall curves from any dataset. Thus, it is unclear whether their dual-process model can properly account for both the monotonic and non-monotonic probability of first recall functions observed in the data.

Usher et al. (in press) present further evidence for sigmoidal recency from a paired-associate memory experiment reported by Waugh and Norman (1965). As discussed above, *TCM-A* can readily produce sigmoidal recall probability functions (as distinct from PFR functions). However, the critical data cited as inconsistent with *TCM-A* was the observation of an apparent interaction between presentation rate and the shape of the recency effect. Specifically, slower presentation rates appeared to give rise to better recall of recent items but lower recall of earlier list items (although Waugh and Norman reported that this interaction was not reliable, a reanalysis by Altmann and Schunn (2002) suggests that the interaction was in fact statistically significant). Although *TCM* has not yet been extended to model paired-associate memory, Altmann and Schunn (2002) showed how a version of the ACT-R model with a continuous decay process (rather than abrupt displacement) provided a good fit to the data.

Recency to primacy shift at fast presentation rates

Usher et al. (in press) report an intriguing new result from their laboratory in which participants exhibit a shift from recency to primacy when presentation rates increase from 0.625 to 5.0 items per second. They argue that this result is inconsistent with *TCM-A*.

Whereas the slow presentation-rate condition exhibited a clear exponentially-decreasing recency effect with little primacy, the fast condition exhibited a strong primacy effect as well as a diminished, yet significant, recency effect. Usher et al.

(in press) focus only on the primacy portion of their results, and illustrate how their buffer-like activation mechanism predicts that primacy items will be able to activate in the buffer during the fast presentation condition because there is not competition from other items. However, once active, the primacy items will prevent subsequent list items from activating in the buffer, thus, giving rise to the enhanced primacy effect.

Although there remains much to be learned about this new presentation-rate phenomenon and its relation to benchmark data in free recall, we do not see how these data necessitate processes beyond those already instantiated within *TCM-A*. One consequence of presenting items at a very fast rate is that participants, who possess limited attentional resources (Marois & Ivanoff, 2005; Nieuwenstein & Potter, 2006), are likely to devote less processing resources to later list items. Furthermore, just as in studies of attentional blink (Bowman & Wyble, 2007), which show the largest effects when words are presented at approximately 5.0 per second, later list items are both forward- and backward-masked by the other items, whereas the first item in the list is only backward-masked. Consequently, the effect of fast presentation rates can be readily implemented in *TCM-A* by increasing the strength of the primacy gradient and decreasing the overall learning rate of the model.

Figure 3 shows *TCM-A* simulations of the data presented in Figure 4 of Usher et al. (in press). The parameters used to generate the serial position curve for the slow presentation rate were the same as those used in Sederberg et al. (in press) to simulate the immediate free recall condition of Howard and Kahana (1999) except for a change in the scale and decay of the primacy gradient ($\phi_s = 1.60$ and $\phi_d = 1.40$) and the standard deviation in the noise of the accumulator ($\sigma = 0.28$). The only parameters changed from the slow condition to simulate the fast presentation rate condition were the scale and decay of the primacy gradient ($\phi_s = 11.18$ and $\phi_d = 1.80$) and the relative weight of pre-experimental to experimental item-to-context associations ($\gamma_{FT} = 0.004$ and

$\gamma_{TF} = 0.400$), which simulates the decreased learning rate as in the amnesia simulations of Sederberg et al. (in press).

As shown in Figure 3, there is a clear increase in the probability of recall of primacy items in the fast presentation rate condition, and a diminished, yet significant, recency effect. Notably, while Usher et al. (in press) demonstrate that their model illustrates the failure to encode subsequent items after the primacy items, they do not illustrate the ability of their model to account for the recency effect apparent in the data simultaneously with the enhanced primacy effect (see also Figures 15, C2, and C3 Davelaar et al., 2005).

Summary

We have shown that contrary to claims made by Usher et al. (in press), the exponential, item-driven context drift process in *TCM-A* can easily support list discrimination and indeed can be shown to fit data on the distribution of prior list intrusions coming from lists of varying lag (see Table 1, Figure 1). The exponential nature of the context drift process in *TCM-A* is perfectly consistent with the finding that recency can be virtually eliminated under some conditions in delayed free recall. Under these conditions one can still obtain a robust contiguity effect (Howard et al., 2006). Although *TCM-A* predicts that the contiguity effect would be somewhat reduced in continual distractor free recall with long distractor intervals, the effect is surely not eliminated (see Figure 2). This prediction is consistent with all extant data. Usher et al. (in press) are correct, however, in noting that for the parameterization reported in Sederberg et al. (in press), the recency and contiguity effects are more sharply tuned than in the data, and the same is true of our simulations of amnesia. Although *TCM-A*'s list length predictions appear inconsistent with Greene's (1986) study, his methods (free recall of paired-associates) were quite different from those simulated in our study (free recall of individual items), so the jury is still out on this point. We have shown that *TCM-A* is consistent with the vast majority of published probability of first recall functions; the three exceptions could be due to rehearsal or other strategic factors. Finally, we have shown that *TCM-A* can fit the qualitative pattern of results in the unpublished study presented by Usher and colleagues as key evidence against our model (see Figure 3).

TCM *and the STS-LTS distinction*

Context in *TCM* has a specific time scale determined by the drift rate in the model. The nature of the contextual drift process produces the recency effect in free recall. This aspect of *TCM* is shared with earlier formulations of contextual drift (e.g., Bower, 1972; Mensink & Raaijmakers, 1988; Murdock, 1997). Although *TCM* predicts recency and contiguity at both short and long time scales, the model does not predict true time-scale invariance (Howard, 2004). However, the approximate time-scale invariance predicted by *TCM* is in line with the data. Because context in *TCM* is not driven by noise but rather by items themselves it is tempting to draw parallels between the contents of context and a short-term store. This point was discussed by both Howard and Kahana (2002) and Sederberg et al. (in press).

Usher and colleagues suggest that the key difference between *TCM* and dual store models is the exponential versus abrupt decay of information. However, this characterization is misleading. Classic dual store (buffer) models are dual process models in that they assume separate retrieval processes govern retrieval from STS and LTS (e.g., Raaijmakers & Shiffrin, 1980, 1981). This is also true of the simulations presented by Davelaar et al. (2005) in their sophisticated context-based dual-store model. Davelaar et al. (2005) show how one could develop a single-retrieval process version of their model (see Appendix C), but they never fully develop this model and show that it can account for short-term and long-term recency and contiguity. In *TCM* a single process governs retrieval of both recent and remote memories. More importantly, items in *TCM* are never retrieved from any kind of buffer (with either exponential or abrupt dynamics). That is, context is not a buffer for storing and retrieving items, and items can never be read out of context. Rather, context maintains a composite representation of recent experiences which is constantly updated and used to cue a weight based association memory system.

In motivating their context-based dual-store model, Davelaar et al. (2005) claimed that *TCM*'s lack of a short-term buffer prevented it from accounting for dissociations between immediate and continual-distractor free recall. Yet, in their comment, Usher et al claim that *TCM* does have an (exponential) short-term buffer, and that this mechanism is responsible for its successful account of the dissociations, as reported in Sederberg et al. (in press). We have always been upfront about the nature of context in *TCM* as being a recency weighted function of item representations (see Howard & Kahana, 2002, p 291). Although *TCM* embodies the distinction between activation and associative structures in memory, we dispute the claim that *TCM* has much more in common with classic buffer models of free recall. Indeed, the essential feature of *TCM* is that item-to-context associations and return context-to-item associations drive the recall process at both short and long time scales, leading the model to offer a parsimonious account of both recency and contiguity effects.

According to both the search of associative memory (SAM) model of Shiffrin and colleagues, and the dual-store model used by Davelaar et al. (2005) to simulate recency effects in immediate and continual-distractor free recall, immediate recency reflects the reporting of items active in STS. In *TCM*, items are never retrieved directly from context. Rather, context serves as the retrieval cue for those items associated with similar context representations. Retrieval of an item then modifies the content of context, which in turn serves as a cue for the next retrieval. This single retrieval process underlies both early and later recalls in *TCM*.

Unlike *TCM*, which forms interitem associations via bidirectional weights between items and context, the weight-based learning mechanism in the Davelaar et al. (2005) model has only a unidirectional connection from context to items, giving rise to no causal link between the recall of one item and the retrieval cue for the next. Instead, the equation governing the activation of items in the buffer contains a term that allows semantically-related items to excite each other while they are both active. Consequently, it may be possible to store direct item-to-item associations if one were to allow for encoding-related changes in the mutual excitation between items in the buffer (Davelaar et al., 2005) and if one provided a means of storing those associations in connection weights. However, it is unlikely that such a process would enable the model to give rise to the long-range contiguity effects seen in the data.

TCM's explanation of the contiguity effect as resulting from an item's retrieval of its temporal context, which in turn serves as a retrieval cue for neighboring items, has been a source of considerable explanatory power (e.g,. Davis, Geller, Rizzuto, & Kahana, 2008; Howard et al., 2007; Howard, Jing, Rao, Provyn, & Datey, Submitted; Polyn & Kahana, 2008; Schwartz, Howard, Jing, & Kahana, 2005; Zaromb et al., 2006). The form of the contextual retrieval process in *TCM-A* also provides a principled account of the temporal asymmetry widely observed in recall tasks (Kahana, Howard, & Polyn, 2008). Moreover, the gradual loss of information in temporal context has enabled *TCM-A* to provide a parsimonious account of recency and contiguity across time scales.

Conclusion

Although the distinction between short- and long-term memory is canonized in our textbooks, portrayed in the descriptions of memory in the media and lay-publications, and is further reified in the clinical literature on memory, we believe that the study of memory has now gone beyond the enumeration and classification of memory systems. A far more promising approach involves the development and testing of models that explicitly describe the complex interaction among the processes that underlie memory function. These processes may involve multiple knowledge representations, encoding processes, and retrieval processes, each with its own temporal dynamics. The challenge for students of memory is to develop parsimonious models that can accommodate the complex functional relations between experimental factors and memory performance as observed in the laboratory.

The temporal context model offers novel insights into the operation of episodic memory by placing the explanatory burden on the interaction between a gradually evolving context representation and weight-based item-to-context associative representations. Although *TCM* encompasses the distinction between activation-based and weight-based memory representations, it makes no distinction between retrieval processes supporting recent and remote memories. This latter distinction, which has long been a cornerstone of dual-process models, is not needed to explain many of the core phenomena concerning memory search in free recall, including those phenomena that had long been a major source of evidence for the short-term long-term memory distinction.

References

Altmann, E. M., & Schunn, C. D. (2002). Integrating decay and interference: A new look at an old interaction. In *Proceedings of the 24th annual conference of the cognitive science society.* Mahwah, NJ: Erlbaum.

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review*, *79*(2), 97-123.

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* (Vol. 2, p. 89-105). New York: Academic Press.

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.

Bower, G. H. (1972). Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (p. 85-121). New York: John Wiley and Sons.

Bowman, H., & Wyble, B. (2007, Jan). The simultaneous type, serial token model of temporal attention and working memory. *Psychological Review*, *114*(1), 38–70.

Brodie, D. A., & Murdock, B. B. (1977). Effects of presentation time on nominal and functional serial position curves in free recall. *Journal of Verbal Learning and Verbal Behavior*, *16*, 185-200.

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*, 3–42.

Davis, O. C., Geller, A. S., Rizzuto, D. S., & Kahana, M. J. (2008). Temporal associative processes revealed by intrusions in paired-associate recall. *Psychonomic Bulletin & Review*, *15*(1), 64-69.

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.

Greene, R. L. (1986). A common basis for recency effects in immediate and delayed recall. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *12*, 413-418.

Hogan, R. M. (1975). Interitem encoding and directed search in free recall. *Memory & Cognition*, *3*, 197-209.

Howard, M. W. (2004). Scaling behavior in the temporal context model. *Journal of Mathematical Psychology*, *48*, 230–238.

Howard, M. W., Jing, B., Rao, V. A., Provyn, J. P., & Datey, A. V. (Submitted). Bridging the gap: Transitive associations between items presented in similar temporal contexts.

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*, 923–941.

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

Howard, M. W., Vankatadass, V., Norman, K. A., & Kahana, M. J. (2007). Associative processes in immediate recency. *Memory & Cognition*, *35*(7), 1700–1711.

Howard, M. W., Wingfield, A., & Kahana, M. J. (2006). Modeling the age-related associative deficit in self-initiated recall. *Psychonomic Bulletin & Review*, *13*(3), 439–445.

Jang, Y., & Huber, D. (2008). Context retrieval and context change in free recall: Recalling from long-term memory drives list isolation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 112–127.

Kahana, M. J., Dolan, E. D., Sauder, C. L., & Wingfield, A. (2005). Intrusions in episodic recall: Age differences in editing of overt responses. *Journal of Gerontology: Psychological Sciences*, *60*, 92–97.

Kahana, M. J., Howard, M. W., & Polyn, S. M. (2008). Associative retrieval processes in episodic memory. In H. L. Roediger III (Ed.), *Cognitive psychology of memory. Vol. 2 of Learning and memory: A comprehensive reference, 4 vols. (J. Byrne, Editor).* Oxford: Elsevier.

Kahana, M. J., Howard, M. W., Zaromb, F., & Wingfield, A. (2002). Age dissociates recency and lag recency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 530–540.

Laming, D. (1999). Testing the idea of distinct storage mechanisms in memory. *International Journal of Psychology*, *34*, 419-426.

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, *9*(6), 296–305.

Mensink, G.-J. M., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review*, *95*, 434-455.

Murdock, B. B. (1962). The serial position effect of free recall. *Journal of Experimental Psychology*, *64*, 482-488.

Murdock, B. B. (1974). *Human memory: Theory and data.* Potomac, MD: Lawrence Erlbaum and Associates.

Murdock, B. B. (1997). Context and mediators in a theory of distributed associative memory (TODAM2). *Psychological Review*, *104*, 839-862.

Murdock, B. B., & Okada, R. (1970). Interresponse times in single- trial free recall. *Journal of Verbal Learning and Verbal Behavior*, *86*, 263-267.

Murdock, B. B., & Walker, K. D. (1969). Modality effects in free recall. *Journal of Verbal Learning and Verbal Behavior*, *8*, 665-676.

Nieuwenstein, M. R., & Potter, M. C. (2006, Jun). Temporal limits of selection and memory encoding: A comparison of whole versus partial report in rapid serial visual presentation. *Psychol Sci*, *17*(6), 471–475.

Polyn, S. M., & Kahana, M. J. (2008). Memory search and the neural representation of context. *Trends in Cognitive Sciences*, *12*, 24–30.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (submitted). Episodic and semantic organization during free recall: The control of memory search. *Submitted*.

Postman, L., & Phillips, L. W. (1965). Short-term temporal changes in free recall. *Quarterly Journal of Experimental Psychology*, *17*, 132-138.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search of associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 14, p. 207-262). New York: Academic Press.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 93-134.

Schwartz, G., Howard, M. W., Jing, B., & Kahana, M. J. (2005). Shadows of the past: Temporal retrieval effects in recognition memory. *Psychological Science*, *16*, 898-904.

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

Shiffrin, R. M. (1970). Forgetting: Trace erosion or retrieval failure? *Science*, *168*, 1601-1603.

Sirotin, Y. B., Kimball, D. R., & Kahana, M. J. (2005). Going beyond a single list: Modeling the effects of prior experience on episodic free recall. *Psychonomic Bulletin & Review*, *12*(5), 787-805.

Usher, M., Davelaar, E. J., Haarmann, H. J., & Goshen-Gottstein, Y. (in press). Short-term memory after all: Comment on Sederberg, Howard and Kahana (2008). *Psychological Review*.

Ward, G., & Tan, L. (2004). The effect of the length of to-be-remembered lists and intervening lists on free recall: A reexamination using overt rehearsal. *Journal of Experimental Psychology Learning Memory and Cognition*, *30*(6), 15.

Waugh, N. C., & Norman, D. (1965). Primary memory. *Psychological Review*, *72*, 89-104.

Zaromb, F. M., Howard, M. W., Dolan, E. D., Sirotin, Y. B., Tully, M., Wingfield, A., et al. (2006). Temporal associations and prior-list intrusions in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 792–804.

Author Note

M.J.K and P.B.S. contributed equally to this article; the order of their authorship is alphabetical. We would like to thank Kenneth Norman, Sean Polyn, and Jonathan Miller for helpful feedback on this manuscript. The authors acknowledge support from National Institutes of Health research grants MH055587 and MH061975 to MJK, MH069938 to MWH, and MH080526 to PBS. Correspondence concerning this article may be addressed to either M. J. Kahana (`kahana@psych.upenn.edu`) or P. B. Sederberg (`persed@princeton.edu`).

Notes

[1] This generalization of *TCM-A* is described in detail in Polyn et al. (submitted) which may be obtained at `http://memory.psych.upenn.edu`. Polyn et al's *Context Maintenance and Retrieval* model, which includes *TCM-A* as a special case, is also able to account for detailed data on temporal, semantic, and source clustering, and their interactions.

[2] The parameters used to simulate Postman and Phillips (1965) differed from those used in fitting data from Howard & Kahana (1999). Specifically, we increased the scale and decay of the primacy gradient ($\phi_s = 3.85$ and $\phi_d = 2.50$), decreased the contextual drift rate due to items ($\beta = 0.56$), increased the drift rate due to distractors ($\beta_{dist} = 0.99$), and decreased the noise in the accumulators ($\sigma = 0.28$).

Number of	Data	Model
Correct responses	5.58 (0.116)	5.38 (0.045)
XLIs	0.150 (0.017)	0.144 (0.031)
PLIs (1 list back)	0.102 (0.014)	0.069 (0.017)
PLIs (2 lists back)	0.026 (0.006)	0.014 (0.005)
PLIs (3 lists back)	0.017 (0.004)	0.006 (0.003)

Table 1

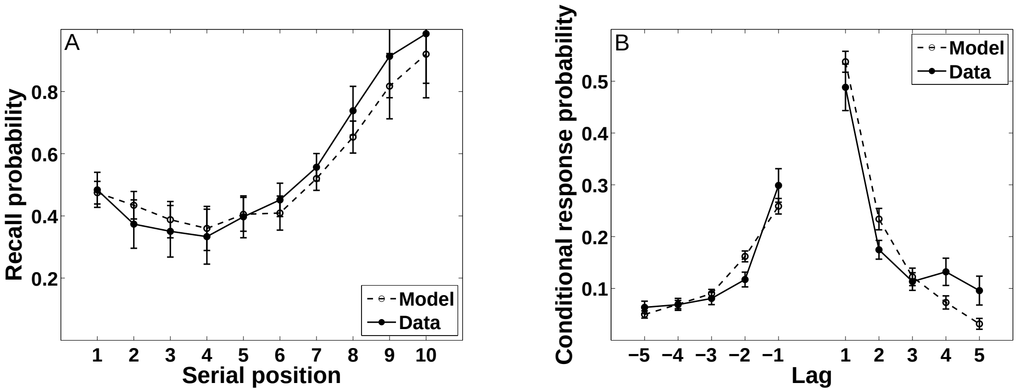
TCM *accounts for the number of correct responses, extra list intrusions (XLIs), and prior list intrusions (PLIs), as well as the distribution of prior list intrusions as a function of list lag. Data are from Kahana et al. (2002). SEMs are shown in parenthesis.*

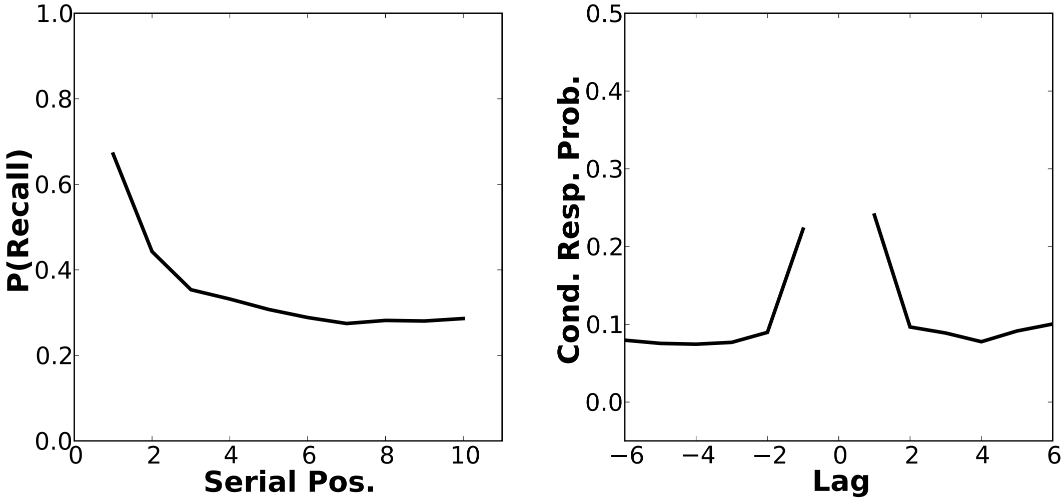
Figure Captions

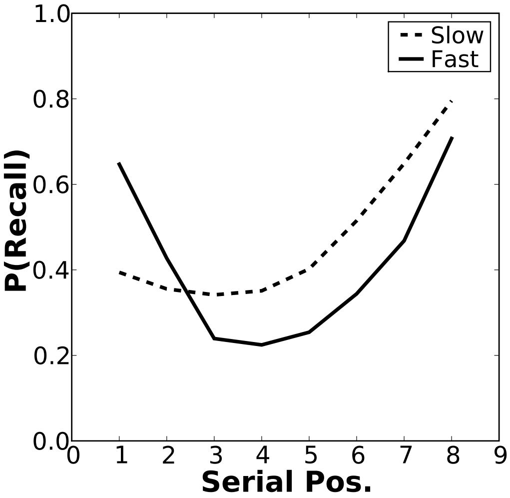
Figure 1. Serial position and contiguity effects in immediate free recall. Using the same model parameters as in Table 1, we can account for the serial position effect (Panel A) and contiguity effect (Panel B) reported in Kahana et al. (2002). Error bars represent ±1 SEM.

Figure 2. Persistence of contiguity without recency. TCM-A simulations of the probability of recall as a function of serial position from the 30-second delayed free recall condition reported in Figure 1 of Postman & Phillips (1965) (left). *TCM-A* simulations of the conditional response probability in continual-distractor free recall with the identical parameters used to eliminate recency in delayed free recall (right).

Figure 3. Shift to primacy with fast presentation rates. TCM-A simulations of the probability of recall as a function of serial position for the fast and slow presentation rate conditions reported in Figure 4 of Usher et al. (in press).


A
Recall probability
0.2
0.4
0.6
0.8
1
2
3
4
5
6
7
8
9
10
Serial position
Model
Data
B
Conditional response probability
0.1
0.2
0.3
0.4
0.5
−5
−4
−3
−2
−1
1
2
3
4
5
Lag
Model
Data


1.0
0.8
0.6
0.4
0.2
0.0
P(Recall)
0
2
4
6
8
10
Serial Pos.
0.5
0.4
0.3
0.2
0.1
0.0
Cond. Resp. Prob.
−6
−4
−2
0
2
4
6
Lag


P(Recall)
Serial Pos.
Slow
Fast
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
6
7
8
9