Psychological Review
2008, Vol. 115, No. 4, 1108–1118

0033-295X/08/$12.00 DOI: 10.1037/a0013725

Short-Term Memory After All: Comment on Sederberg, Howard, and Kahana (2008)

Marius Usher and Eddy J. Davelaar
University of London

Henk J. Haarmann
University of Maryland

Yonatan Goshen-Gottstein
Tel Aviv University

P. B. Sederberg, M. W. Howard, and M. J. Kahana (2008) have proposed an updated version of the temporal-context model (TCM-A). In doing so, they accepted the challenge of developing a single-store model to account for the dissociations between short- and long-term recency effects that were reviewed by E. J. Davelaar, Y. Goshen-Gottstein, A. Ashkenazi, H. J. Haarmann, and M. Usher (2005). In this commentary, the authors argue that the success of TCM-A in addressing the dissociations is dependent not only on an episodic encoding matrix but—critically—also on its implicit use of a short-term memory store—albeit exponential rather than buffer-like. The authors also highlight some difficulties of TCM-A in accounting for these dissociations, and they argue that TCM-A fails to account for critical data—the presentation-rate effect—that dissociates exponential and buffer-like models.

Keywords: free recall, short-term memory, long-term recency, temporal context

At the end of the 19th century, William James (1890) drew the distinction between primary and secondary memory in terms of the content of memory that is part of the conscious-present and is subject to capacity limitations versus that which is not. This distinction was later mapped onto short-term memory (STM) and long-term memory, which fits naturally within the Hebbian framework of neural-reverberation versus structural-synaptic changes (Hebb, 1949). Moreover, this distinction is central for the dual-memory model (Atkinson & Shiffrin, 1968; Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Raaijmakers & Shiffrin, 1981; Waugh & Norman, 1965), which assumes an STM store or buffer, corresponding to activated representations in a heightened state of accessibility and prone to displacement by incoming information (Glanzer, Gianutsos, & Dubin, 1969), in addition to a long-lasting episodic encoding. This dual model has been successful in accounting for vast amounts of cognitive and neuropsychological data (in particular, the recency effect in free-recall type paradigms, partially reviewed by Sederberg, Howard, & Kahana, 2008), and an STM model based on neural reverberations was recently proposed to account for data in serial order tasks (Botvinick & Plaut, 2006).

During the 1980s, however, the basis for this distinction was challenged by memory theorists who, in response to the discovery of long-term recency (see review in Sederberg et al., 2008), argued that STM, even in immediate free recall, is a redundant concept (Crowder, 1982, 1993; Greene, 1986b). Accordingly, it was argued that one should abandon the concept of a short-term store in favor of a single-store account with a retrieval mechanism sensitive to the temporal discriminability (or recency) of the memoranda on all timescales (Brown, Neath, & Chater, 2007; Crowder, 1982; Glenberg, Bradley, Kraus, & Renzaglia, 1983). The impact of the temporal-discriminability account can be seen in psychology textbooks in which STM is presented under headings such as *The Rise and Fall of Short-Term Memory*, even though the same textbooks cover working memory and capacity limitations. This confusing situation reveals that the field has not yet settled on an agreed understanding of the topic. Recently, we presented and reviewed evidence for dissociations between short- and long-term recency. Moreover, we presented a model that accounts for these dissociations under the assumption that although a mechanism of retrieval from long-term episodic memory (sensitive to recency) is indeed necessary for accounting for much of the recency data, an activation buffer must also be postulated if one is to account for the full range of dissociations between short- and long-term recency effects (Davelaar et al., 2005). In their article, Sederberg et al. (2008) took the challenge of examining whether a memory system without an STM component can account for the same dissociations on the basis of temporal discriminability alone (implemented via contextual change), as a general property of memory.

To this end, Sederberg et al. (2008) proposed a new version of the temporal-context-model (TCM; Howard & Kahana, 2002), labeled TCM-A, that reproduces some patterns found in some of

Marius Usher and Eddy J. Davelaar, School of Psychology, Birkbeck College, University of London, London, United Kingdom; Henk J. Haarmann, Center for Advanced Study of Language, University of Maryland; Yonatan Goshen-Gottstein, Department of Psychology, Tel Aviv University, Tel Aviv, Israel.

Yonatan Goshen-Gottstein is supported by the Israel Science Foundation (Grant 894-01). We thank Mike Kahana and Ben Murdock for making the data sets publicly available.

Correspondence concerning this article should be addressed to Marius Usher, School of Psychology, Birkbeck College, University of London, Malet Street, London, United Kingdom WC1E 7HX. E-mail: m.usher@bbk.ac.uk

the dissociations we previously reviewed. Like its TCM precursor, this model can account for recency functions and contiguity effects across timescales. We find this contribution of great interest, first for offering a rigorous process model of memory in free recall, and second (and perhaps more importantly) for opening a window onto its powers and limitations. In this commentary, we start with a discussion of the commonalities between Sederberg et al.'s TCM-A and our dual-component memory model. By examining relevant data, we then discuss the differences between the models and the implications of these differences for the debate on STM. To anticipate, we believe that Sederberg et al.'s article does very little to demonstrate that STM is a redundant concept, and, if at all, it makes the need for an STM store even more obvious.

Common Characteristics: An STM Component in TCM-A

Central to the ability of TCM-A to account for dissociations between short- and long-term recency is the distinction between the state of context and the matrix of learned (experimental) connections between the context and item representations. What in TCM-A is called context is a layer of units (separate from the item layer) that is activated by the item layer (in a one-to-one fashion) and is subject to a type of decay or competition, which keeps the state normalized. As admitted by Sederberg et al. (2008; see their *Short-Term Memory and Temporal Context* section, p. 908), the state of context can be simply expressed as a weighted average of patterns that are associated with the items presented for recall, with exponentially decaying weights. In other words, the state of context is a trace in which the activation of early items decays by a constant factor, such that the addition of each new encoded item leads to further decay of the previous items. Thus, for all purposes, the state of the TCM-A context is a type of STM store whose role is twofold. First, like in the search of associative memory model (SAM; Raaijmakers & Shiffrin, 1981), it mediates indirect associations between the items. Second, it is used as a memory cue. Because the state of context (which corresponds to the recently processed items) decays after each distractor interval, it results in a weaker cue in delayed free recall (DFR) and continuous-distractor free recall (CDR) than it does in immediate free recall (IFR), as is the case in the dual-store models. We obviously do not argue that the nature of the memory trace in TCM-A context is the same as that in SAM and in our model (see section Differences Between TCM-A and Dual-Store Models). What we do argue is that it is this STM inheritance—the distinction between the episodic matrix and the decaying trace of context (following each new item or distractor interval)—that enables the TCM-A to capture some of the fundamental dissociations between short-term and long-term recency.

For example, it is the decay of the trace after each distractor that explains the slower response latencies in the first recall in DFR and CDR as opposed to in IFR. Likewise, this decay mediates the intact performance for the last list item for amnesia patients in IFR but not in the CDR task. Although in TCM-A the context state and episodic matrix are continuously combined to drive recall (see Figure B1 in Appendix B of Davelaar et al., 2005, for a similar process during recall), one can still make an important distinction between the specific contributions of context state and those of the long-term memory episodic matrices. Whereas a distractor interval at the end of a list makes the context state more flat (Figure 11 in Sederberg et al., 2008) and the recall driven by this input and the episodic matrix has a diminished recency, the opposite is the case in amnesia. In amnesia, the episodic matrix is deficient, but the state of context (i.e., the exponentially decaying weighted activity of presented items) is intact and can drive the recall of recency items in IFR. Thus, although amnesiacs have a deficit at all serial positions in CDR (due to the lack of learned connections), they have an intact recency at the last item in IFR (due the intact context state). This is precisely the mechanism by which dual-store models account for the amnesiac deficit.

In other words, although portrayed as a single-store model, TCM-A is also a dual-component model, with the two components (the state of context, which consists of a decaying state of the actual items and the episodic matrix) continuously interacting. This underscores the need for a dual-store architecture to capture a wide range of free-recall data. Thus, we argue that the redundancy of the dual-model framework, which, following Hebb (1949), distinguishes between activation and structural components to memory, has not in any way been established. Rather, it appears that the fundamental notion behind an STM store, that of a decaying item representation, is still at the core of a model that has been proposed specifically with the aim of demonstrating its redundancy.

Another common feature in TCM-A and our model is that neither satisfies the principle of temporal-scale invariance, according to which forgetting functions maintain their form when the presentation of material is scaled in time, such as in the transition from IFR to CDR. This is not a criticism of TCM-A as we do not think that time invariance applies in IFR paradigms other than as a metaphor. However, we believe it is important to clarify this because TCM is sometimes associated with the principle of time-scale invariance (see section *Approximate Scale Invariance* in Howard & Kahana, 2002) and could be thought of as implementing Crowder's (1976) principle of temporal distinctiveness. The fact that TCM-A is not time invariant (see Howard, 2004) can be seen in Figure 11 of Sederberg et al. (2008), where the input to the choice units at the time of the first recall is very different in IFR (extended recency) and CDR (very narrow recency), resulting in very different recency functions in both first and total recall. This is a direct result of the contextual change mechanism of TCM-A, resulting in a much steeper gradient of items in the state of context in CDR as compared to IFR. When a single distractor interval at the end of list is inserted (as in DFR), the relevant context state is simply scaled down by a multiplicative factor. When distractor intervals are inserted between each list items, the gradient becomes steeper, and no scaling (such as the C_V factor in Equation 10 of Sederberg et al., 2008) can restore the gradient observed in IFR. The question of whether pure temporal-discriminability models (Brown et al., 2007) can account for dissociations between IFR and CDR and other STM data is an important one, which we defer to the Discussion section.

Differences Between TCM-A and Dual-Store Models

There are several differences between TCM-A and dual-store models. One difference is the fact that TCM-A (but not our dual-component model) assumes that after the recall of each item, the state of the context (or STM store) is updated such that it restores part of the item's context, and this new context state is used to cue memory recall. This is an important feature in all TCMs, which helps them to account for contiguity effects in both IFR and CDR. Although we have not included such an assumption in our model, we do not view it as inconsistent with it (see Elhalal, 2008, for a dual-component model with a TCM-type of context). We will focus here on two more important differences. The first is the fact that TCM-A lacks a list context, and the second is the nature of the STM store (or state of context in TCM-A).

First, although TCM-A (as well as its precursors) emphasizes fluctuating context as a mechanism for episodic-memory encoding, its context is a function of previously activated items and is, therefore, not independent of them (unlike other context fluctuations models; Estes, 1955; Mensink & Raaijmakers, 1988). This is important because it is precisely the assumption of a list-context representation independent of item representations, which was the key mechanism traditionally used in memory models, such as SAM, to discriminate list information for a given item (Anderson & Bower, 1972; Mensink & Raaijmakers, 1988). Because no such independent context representation exists in TCM-A, the model (in its present form) lacks a flexible mechanism enabling it to distinguish items from one list from items of previous lists. In the absence of list context, the mode of operation of TCM-A is purely associative, driven by associations between items encoded during the list presentation and mediated via the context layer; functionally, TCM-A is equivalent to a modified buffer component (see below) of SAM, which encodes associations between items while they are coactive in the buffer. As a result of the lack of list context, TCM-A can only be applied to one-list paradigms so it cannot truly account for phenomena that require multiple lists, such as the novel proactive interference data that were presented in Davelaar et al. (2005; see *Wrong Prediction for List-Length Manipulation in CDR* section below). It remains to be seen how TCM-A may be extended to include list context and applied to multilist paradigms.

Second, we agree that the major difference between TCM-A and dual-store models (such as SAM or ours) is not the existence of activation versus weight-based processes (as this takes place in both) but rather the nature and dynamics of the activation component. Unlike traditional STM stores (e.g., SAM), which have a fixed number (three–four) of slots from which items are displaced when new information is presented, the TCM-A context has gradual activations, which enable items to decay by a multiplicative factor when a new list item is presented or when a distractor interval has occurred. Our activation buffer (Davelaar et al., 2005; Usher & Cohen, 1999) is in an intermediate position. Like TCM-A, the activation's state is gradual, which is crucial for accounting for the precise shape of serial-position functions and their dependency on presentation rates (see next section). But unlike TCM-A, its activation dynamics approximate a displacement process caused by mutual inhibition between cell assemblies that are subject to nonlinear attractor dynamics, making it functionally similar to traditional buffer models (see Figure 1). This difference in activation dynamics, which we label for short as *exponential versus abrupt dynamics* (or exponential versus buffer-like activation store), has important consequences for the way the models work. For example, it is the exponential nature of its activation store that allows TCM-A to encode the contiguity between list items in the CDR task (as some item activation survives in the context state after a distractor interval) without having to assume a separate episodic system (Davelaar et al., 2005). However, although both models account for broad qualitative patterns of the data in free-recall tasks, they diverge on quantitative details and on key additional data. In the following section, we compare the models' accounts on the dissociations between short- and long-term recency before turning to some other relevant data.

Divergences in Accounting for Recency Functions

Although TCM-A provides a broad qualitative account for recency data in free-recall paradigms, we believe that, at present, its account does not capture the data as well as our dual-store model does. In the following, we list four data patterns in which TCM-A does not provide an accurate account of recency data. We finish with a fifth pattern, that of the list-length effect, where the model's

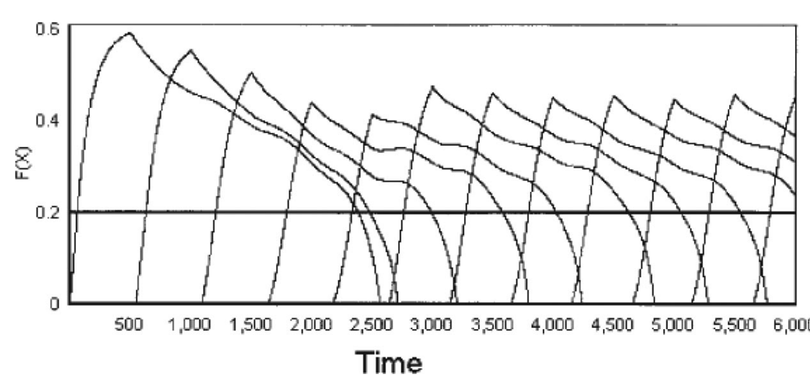

Figure 1. Activation dynamics in the activation buffer. The activation, $F(x)$, of each representation is plotted against the time (in model iterations). Each of the 12 sequentially activated representations received external input for 500 iterations only. The above-threshold activation of a representation when its external input is taken away forms part of the content of short-term memory, which is limited to about four items (count the number of activation profiles that are above the threshold at any time). From "The Demise of Short-Term Memory Revisited: Empirical and Computational Investigations of Recency Effects," by E. J. Davelaar, Y. Goshen-Gottstein, A. Ashkenazi, H. J. Haarmann, and M. Usher, 2005, *Psychological Review, 112,* p. 11. Copyright 2005 by the American Psychological Association.

predictions are categorically wrong. For all the results, we point to the source in the model's mechanism that is responsible for the differences between the data and the model's simulations.

1. Recency in DFR

The complete elimination of the recency effect in DFR was the major motivating factor for assuming an abrupt, displacement-based buffer store in the first place. Consistent with the data reported by Howard & Kahana (1999), TCM-A predicts an attenuation of recency in DFR rather than its complete elimination as a result of its exponential activation store (Figures 3, 4, and 11 of Sederberg et al., 2008). The problem is that most of the DFR data that motivate the dual-store model (e.g., Glanzer & Cunitz, 1966; Haarmann & Usher, 2001; Postman & Philips, 1965) show a complete elimination of the recency effect. TCM-A could further reduce or perhaps eliminate recency in DFR by increasing the amount of contextual change during distractor intervals. Such a change, however, should result in diminished contiguity effects in CDR. It is an open question whether such reduced contiguity effects will be obtained in a CDR task with materials (word lists and distractor intervals) that result in null DFR recency. The problem is that if this is the case (and unless the contiguity effect will still be significantly higher than 0), single-store models like TCM are not the only models consistent with the result; SAM or other explicitly dual-store models can account not only for contiguity effects in CDR (as a result of residual buffer information) for distractor intervals that have an attenuated DFR recency but also for the lack of contiguity with distractor intervals that eliminate DFR recency.

2. Only One-Item Recency in CDR

In order to attenuate recency in DFR, TCM-A uses a relatively large value for the β_{dist} parameter (1.5 times larger than β during item presentation). Because this decay takes place after every item in CDR and because the absolute values of activations, not only their relative values, affect retrieval, this yields only a one-item recency effect (Figures 3, 4, 8, and 11 of Sederberg et al., 2008). However, the empirical data show a more extended recency gradient in CDR. This is because the information of all previous items (other than the last one) in the context state at the end of list presentation has decayed to levels that are buried deep within the noise.

3. Only One-Item Lag Recency in Conditional Response Probability in CDR

This outcome (Figure 5 of Sederberg et al., 2008) is due to the same mechanism discussed above, that is, the fast decay of context (high β_{dist}) during a distractor interval.

4. Only One-Item Intact Recency in IFR for Amnesia

As a result of its exponential buffer, TCM-A predicts that recall is intact for only the last list item in IFR for amnesia patients (Figure 9 of Sederberg et al., 2008). This is at odds with the experimental evidence (Baddeley & Warrington, 1970; Carlesimo, Marfia, Loasses, & Caltagirone, 1996) that shows a more extended range of intact recency of about three–four items (intact recency slope) and an intact digit span. Due to its buffer-like property, our activation store (Davelaar et al., 2005) predicts intact performance on the last three–four items.

5. Wrong Prediction for List-Length Manipulation in CD

Sederberg et al. (2008) use a list-length manipulation as a substitute for across-lists proactive interference (PI). In doing so, TCM-A predicts that increasing list length affects the recency items more in CDR than it does in IFR (Figure 8 of Sederberg et al., 2008). We note first that, even if one accepts that manipulating list length is equivalent to inducing PI, the TCM-A simulations do not capture the qualitative pattern of the data. This is because in the data—but not in the model (Figure 8 of Sederberg et al., 2008)—the PI effect in CDR is uniform across all serial positions, whereas TCM-A predicts a larger effect of PI on the last serial positions compared to earlier list positions. Thus, the model predicts an interaction between PI and serial position in CDR rather than an additive effect.

More importantly, the curious thing about the TCM-A prediction that list length dissociates IFR and CDR is that it contradicts empirical data, originally reported to reject the need for an STM store (Greene, 1986a). These data (which were accounted for in our dual-store model, see Figure 2) show that recency in both IFR and CDR is *not* affected by list length. Unlike TCM-A, our model was able to account simultaneously for the association between IFR and CDR with list length and for the dissociation with PI.

The key for the distinction between list length (longer list) and PI (two shorter lists) is the retrieval phase between the two lists

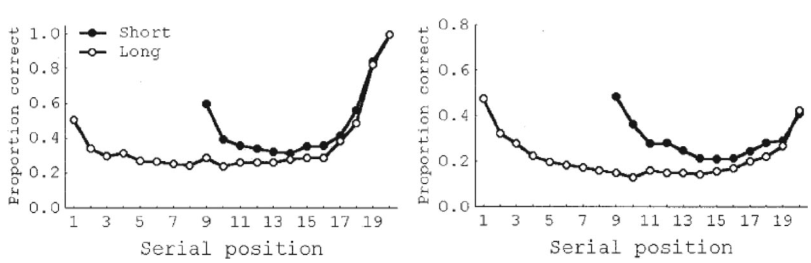

Figure 2. Model predictions for our dual-memory model to account for list-length association between immediate free recall (left panel) and continuous-distractor free recall (right panel). From "The Demise of Short-Term Memory Revisited: Empirical and Computational Investigations of Recency Effects," by E. J. Davelaar, Y. Goshen-Gottstein, A. Ashkenazi, H. J. Haarmann, and M. Usher, 2005, *Psychological Review, 112,* p. 14. Copyright 2005 by the American Psychological Association.

(present in the PI paradigm but absent in longer lists), which resets the context to previous states making the two list contexts more overlapping. Although it is possible that TCM-A can be extended to deal with multiple lists, one may note that at present it makes an incorrect prediction about list length (compare Figure 8 in Sederberg et al., 2008, with our Figure 2 and data in Greene, 1986a).

Decisive Data: Form of Serial Position Function and Presentation Rate Effects

We have examined so far a number of dissociations between recency functions in IFR and in CDR that were discussed by Sederberg et al. (2008). In previous articles (Davelaar et al., 2005; Davelaar, Haarmann, Goshen-Gottstein, & Usher, 2006; Talmi, Grady, Goshen-Gottstein, & Moscovitch, 2005), we addressed additional dissociations (directed output order, final-recall negative recency, semantic effects, and neurofunctional effects) that readers may like to examine further. Furthermore, a novel dissociation has recently been reported between IFR and CDR that poses a difficult challenge to TCM (Farrell, 2008). We believe, however, that the most sensitive data that can disentangle the predictions of a buffer-like versus an exponential activation store (or temporal-discriminability model) involves the examination of the detailed shape of serial position functions and the effects of presentation rate on it.

Consider first the form of the serial position function at recency. In many data sets, in free recall the serial position functions have a sigmoidal shape (Howard & Kahana, 1999; Murdock, 1962): the recall drops abruptly after two–three items from the end of list (see Figures 3A and 3B). Stochastic buffer-like models predict (for low noise levels) such a sigmoid serial position function as a natural outcome of the fact that all items in the activation buffer are reported, and the probability for an item to be in the buffer is sigmoidal with recency (Davelaar et al., 2005); such sigmoidal recency was one of the major motivations for adopting a buffer-like store with a limited number of slots within the modal model (see Figure 18 in Atkinson & Shiffrin, 1968). The key prediction of an exponential buffer, on the other hand, is an exponentially shaped retrieval strength (i.e., the input into the items as a result of the end of list probe; see Figure 11 in Sederberg et al., 2008); similar (J-shaped) retrieval strength functions are predicted by

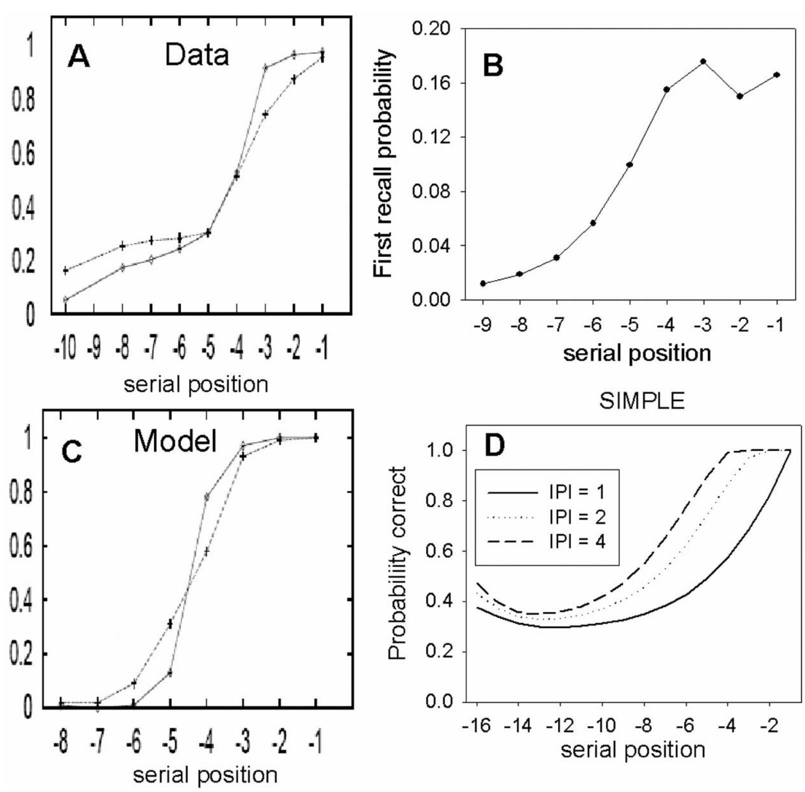

Figure 3. A) Serial position functions for cued recall at fast/slow (steep/shallow sigmoids) presentation rates in cued recall for data from Waugh and Norman (1965). B) First recall probability function showing sigmoidal recency gradient (averaged data from Murdock, 1962 and Murdock & Okada, 1970). C) Serial position functions for cued recall at fast/slow (steep/shallow sigmoids) presentation rates in cued recall for the activation model. D) Predictions of SIMPLE when the interpresentation interval (IPI) is increased by a factor of two or four. Panels A and C are from “Short-Term Memory and Selection Processes in a Frontal-Lobe Model” (p. 84, in *Connectionist Models in Cognitive Neuroscience*, edited by D. Heinke, G. W. Humphries, & A. Olsen), by M. Usher and J. D. Cohen, 1999, New York: Springer-Verlag. Copyright 1999 by Springer-Verlag. Reprinted with permission.

temporal-discriminability models (Brown et al., 2007). If the recall probability is approximately linearly related to retrieval strength, then this results in a similarly shaped recency gradient in cued recall or first-recall probability free recall (see Figure 4 of Sederberg et al., 2008).

Models such as TCM-A can still account for sigmoidal recency functions in IFR (Figure 3 in Sederberg et al., 2008) because, following the first recall, one is likely (due to contiguity effects to retrieve another recency item). Thus, to distinguish between exponential and abrupt buffers, one needs to examine serial positions in paradigms that report a single item: cued recall or first output IFR. Data in cued recall (Waugh & Norman, 1965) shows a clear sigmoidal pattern. First output probability data in IFR are more ambiguous. Whereas some reports show J-shaped recency (Howard & Kahana, 1999), other reports show sigmoidal recency gradients (Murdock, 1962; Murdock & Okada, 1970; see Figure 3B). Both patterns can be obtained in traditional buffer models depending on the noise magnitude. In order to account for such sigmoidal recency functions, exponential buffer models or temporal-discriminability models need to introduce a strength threshold such that all items whose retrieval strength exceeds the threshold (no matter by how much) are equally retrieved (Brown et al., 2007).

The effect of presentation rate on the serial position functions can further differentiate these types of models. In Figure 3C, we show the prediction of the activation buffer on the probability of having an item active at the end of recall (Usher & Cohen, 1999) as a function of its serial position in the list for slow/medium presentation rates, following the cued-recall experiment of Waugh & Norman (1965). We observe that, as in the data, the serial position function has a sigmoidal shape, and the total recall does not differ much as a result of the change in present presentation rate (thus approximating a fixed capacity buffer). However, the slope of the sigmoidal function is steeper at slower, as compared to medium, presentation rates (see Figure 3C). The reason why the activation buffer makes this prediction is its nonlinear dynamics; at slow presentation rates, the duration of the competition is longer per item and favors the more recent items due to their more recent sensory input.

As opposed to buffer-like models, exponential activation stores or temporal-discriminability models, such as SIMPLE (Brown et al., 2007), even with the auxiliary assumption of a retrieval-strength threshold for recall, cannot account for the increase in the slope of the sigmoid without a change in total recall (Figure 3D). Moreover, an even stronger challenge to such models is the drastic effect of very fast presentation rates on the serial position function in probed recall. In our recent article (Davelaar et al., 2005), we examined a surprising prediction of the activation buffer at fast presentation rates (100–200 ms/item). We showed that at such fast rates (where no rehearsal can take place), the model predicts a dramatic shift from recency to primacy so that earlier conditions in a list are reported more than later positions are. This prediction, which is the outcome of the competition between the items in the activation buffer that at short presentation dominates the input advantage to recent items (Figure 4A), was supported by our cued-recall experiments (see Davelaar et al., 2005). Similar results confirming a primacy type of serial position curve have been obtained in our lab at fast presentation rates (200 ms/word) in IFR (Figure 4B). Moreover, very strong primacy-like serial position functions, consistent with these predictions, have recently been reported by Nieuwenstein and Potter (2006) in IFR of six-letter sequences presented at 106 ms/letter.

As both exponential activation store and temporal-discriminability models have recency as their main defining principle, they cannot account for this qualitative change in the form of the serial position function. By contrast, a dual-store model with time-sensitive competition among activated items in short-term memory offers a unitary account for the impact of presentation rate at both primacy and recency positions.

Discussion

We have previously reviewed a number of dissociations between short-term and long-term recency, which were accounted for

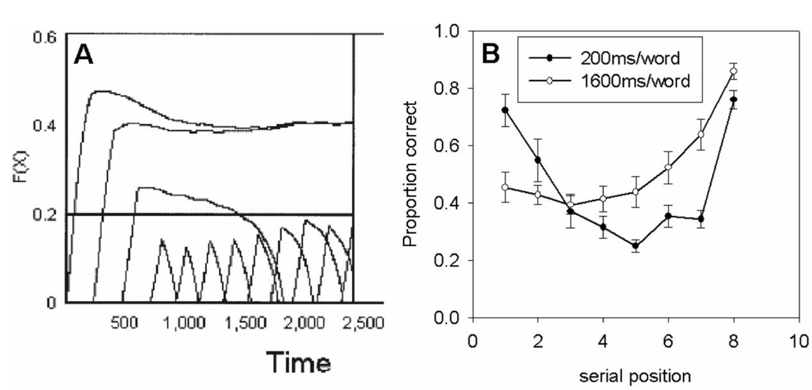

Figure 4. A) Activations in response to fast presentations giving rise to primacy. From "The Demise of Short-Term Memory Revisited: Empirical and Computational Investigations of Recency Effects," by E. J. Davelaar, Y. Goshen-Gottstein, A. Ashkenazi, H. J. Haarmann, and M. Usher, 2005, *Psychological Review, 112,* p. 21. Copyright 2005 by the American Psychological Association. B) Serial position function in immediate free recall of eight-word sequences for fast/slow presentation rate (200 vs. 1,600 ms/word). From *A Closer Look at Presentation Rate Effects in Free Recall*, by D. A. Hawkins and E. J. Davelaar, January 2005, poster presented at the meeting of the Experimental Psychological Society, London, England.

within the dual activation/weight-based model of episodic memory (Davelaar et al., 2005). These dissociations raise a challenge for unitary-memory models, which, following Crowder (1982), reject the reality of—or need for—an activation component. TCM-A is a process model of free recall that, allegedly, follows this unitary-model program and takes up the challenge to account for these dissociations.

The main assumption of TCM-A is that memory for list items is encoded via the associations between list items mediated by a state of context, which, in turn, is a function of the items processed (a weighted exponentially decaying average) and is also used to probe memory retrieval. This model captures broad qualitative patterns in free-recall paradigms and, more importantly, it addresses some dissociations between STM and long-term memory paradigms (IFR/CDR). For example, the model accounts for shorter retrieval latency in IFR as compared to CDR, for the dependency of the conditional-response probability on output order in IFR but not in CDR, and for a smaller deficit on the final list item in IFR compared with CDR that is found with amnesiac patients. Our analysis of the models reveals, however, that all these successes are the outcome of a short-term store, which is implicitly designed into the model. That is, TCM-A makes use of two memory components: a decaying activation trace (the state of context) and an associative matrix between item and context layers.

The divergence between TCM-A and our dual-component model centers on two aspects: the assumption of a list-context representation in our model (and its lack in TCM-A) and the nature of the activation store. For TCM-A, the activation store (or state of context) has the property of an exponentially decaying (with intervening items or distractors) store, which stands in contrast with the more buffer-like activation store assumed in our model. Due to its exponential activation store, TCM-A can account easily for contiguity effects between list items, even when these are separated by distractor intervals. However, although the exponential buffer allows TCM-A to address some of the dissociations discussed, its quantitative account of the data is less good, especially with regard to manipulations such as list-length, PI, and the extended (three–four items) sparing of recency in amnesia. Future research (extending TCM-A) is needed to examine this model's ability to address these dissociations, as well as other recently reported data (on long-lag transitions in the order of recall) that pose a problem for TCMs (Farrell & Lewandowsky, in press).

The most decisive data against models with an exponential activation store (such as the context state in TCM-A) appear to be the sigmoidal form of the recency gradient and the effect of the fast and slow presentation rates on it. Although this adds support against an exponential buffer, we do not think that such buffers have no role in memory (see also Cowan, 2001), as it is possible that multiple types of activation systems exist and work in tandem within various brain systems (e.g., exponential short-term storage in the hippocampus, Howard, Fotedar, Datey, & Hasselmo, 2005, vs. attractor dynamics in prefrontal cortex, Wang, 2001), contributing together to task performance. In the following, we discuss a number of more general implications of the results presented here to the structure of and processes in memory.

The Temporal-Discriminability Principle

Although TCM-A is often perceived as implementing the principle of temporal discriminability explicitly assumed in some models (Brown et al., 2007; Glenberg & Swanson, 1986; Neath & Crowder, 1990), it does not have the properties typically associated with these models. In particular, the model does not satisfy the principle of temporal-scale invariance. One recent temporal-discriminability model (Brown et al., 2007) has strongly advocated the idea that the temporal discriminability between items (subject to Weber law type of representations) is the major principle that accounts for patterns of recall at both short and long timescales.

Although the temporal-discriminability model by Brown et al. (2007) accounts for a vast amount of memory data, it has not been shown to offer a complete account for most of the dissociations between short- and long-term recency. For example, the model does not account for the dissociation in the performance of amnesic patients nor for the dissociation in conditional lag-recency profiles with output order, and it cannot reproduce the immunity to PI in IFR for the last three–four list items (Neath & Brown, 2006). More critical, however, is the fact that due to its central assumption, temporal-discriminability models like TCM-A, cannot account for the effects of presentation rate on the shape of serial position function, and in particular, the increase in the slope of the sigmoidal serial position function and the shift from recency to primacy caused by faster presentation rates.

We do not dispute the fact that retrieval from episodic long-term memory is subject to important regularities, such as temporal discriminability. We, too, are sympathetic to the idea that laws of nature should work at all scales (Brown et al., 2007). And yes, we also aspire to see the study of memory reach the same rigorous standards as in physical sciences. However, we believe the data we have discussed support the presence of an additional component that mediates recency (caused by an activation buffer) on top of the other regularities of episodic memory recall, an idea analogous to the appearance of nuclear forces at short distances in the domain of physical interactions.

The Role of Context

The effect of context in memory encoding and retrieval is one of the most important aspects in memory research (Tulving & Thomson, 1973). Context, in a somehow different sense (task context), is also an important part of attentional control (Cohen, Dunbar, & McClelland, 1990; Cohen & Servan-Schreiber, 1992). Although the state of context plays a major role in the functioning of TCM-A, its interpretation in relation to either of the processes above is unclear. In TCM-A, the context is described as corresponding to a set of features of concepts previously processed. In other words, TCM-A context is defined as "an amalgam of many prior contexts in which the item has appeared" (Sederberg et al., 2008, p. 897) and thus, by activating many of the contexts in which an item appeared, TCM-A is essentially activating the concept itself. This makes the TCM context a type of memory trace, excluding both the task context (environmental factors, e.g., type of room) and the list context and even autonomously generated thoughts, which were the natural interpretation of the random fluctuation context models (Estes, 1955; Mensink & Raaijmakers, 1988). A more recent model within the TCM framework makes some important steps in incorporating task context and list context (Polyn, Norman, & Kahana, 2008). Future work is needed to develop models that address the role of the various types of context and their interactions in memory.

Beyond List Memory

We have focused here on list memory, in particular, free-recall paradigms. Recently, an attractor network was proposed to account for short-term maintenance of information in serial order tasks (Botvinick & Plaut, 2006). Moreover, the activation memory buffer framework is consistent with research outside of field of verbal memory. For example, a capacity-limited buffer at the interface of attention and memory accounts for data in visual STM (Luck & Vogel, 1997), visual tracking (Pylyshyn & Storm, 1988), and individual differences in higher-level cognition (Engle, Tuholski, Laughlin, & Conway, 1999). Such data are a major motivation for the capacity-limited activation state of Cowan's (1995, 2001) embedded processes model. Furthermore, a lexical-semantic buffer has been used to account for data in individual differences in language tasks with normal (Haarmann, Davelaar, & Usher, 2003) and aphasic participants (Martin, Lesch, & Bartha, 1999; Martin, Shelton, & Yaffee, 1994), and a multimodal episodic buffer was recently introduced by Baddeley (2000) to account for sentence recall.

The Relation With Consciousness

It was the relation with consciousness (the conscious present) that motivated James to propose the distinction between primary and secondary memory. In the concluding section of his article, Crowder (1993, p. 145) observed that although he shares James's intuitions about the psychological present, he does "not trust them for one moment," suggesting that they are of the same type that "led to a firm belief in a geocentric universe and a flat earth." Surely, Crowder's skepticism has stimulated the field and resulted in richer regularity patterns than were known before. At the same time, however, we believe that the data that have accumulated within experimental psychology and the natural mapping of this distinction within the Hebbian framework of neural reverberation versus structural synaptic modifications (see, e.g., Warden & Miller, 2007, Figure 8 for neural reverberation associated to multiple items in memory recall) provide a vindication of those widely shared intuitions. Models of the activation store, such as the activation buffer (Davelaar et al., 2005) and the context state of TCM-A, give different insights into the nature of the conscious present. Whereas in the former this is abrupt, in the latter it fades continuously into the background. Other activation stores (Baddeley, 2000; Cowan, 1995; Oberauer, 2002) suggest other insights. Unlike in the 1980s (when the demise of STM was first proposed), the study of consciousness has now become a legitimate (and very productive) topic of scientific investigation. Future work is surely needed to examine the relation between the activation/neural reverberation and the content of consciousness.

References

Anderson, J. R., & Bower, G. H. (1972). Recognition and retrieval processes in free recall. *Psychological Review, 79,* 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, pp. 89–195). New York: Academic Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4,* 417–423.

Baddeley, A. D., & Warrington, E. K. (1970). Amnesia and the distinction between long- and short-term memory. *Journal of Verbal Learning and Verbal Behavior, 9,* 176–189.

Botvinick, M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review, 113,* 201–233.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114,* 539–576.

Carlesimo, G. A., Marfia, G. A., Loasses, A., & Caltagirone, C. (1996). Recency effect in anterograde amnesia: Evidence for distinct memory stores underlying enhanced retrieval of terminal items in immediate and delayed recall paradigms. *Neuropsychologia, 34,* 177–184.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing model of the Stroop effect. *Psychological Review, 97,* 332–361.

Cohen, J. D., & Servan-Schreiber, D. (1992). Context, cortex and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review, 99,* 45–77.

Cowan, N. (1995). *Attention and memory: An integrated framework.* New York: Oxford University Press.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24,* 87–185.

Crowder, R. G. (1976). *Principles of learning and memory.* Hillsdale, NJ: Erlbaum.

Crowder, R. G. (1982). The demise of short-term memory. *Acta Psychologica, 50,* 291–323.

Crowder, R. G. (1993). Short-term memory: Where do we stand? *Memory & Cognition, 21,* 142–145.

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.

Davelaar, E. J., Haarmann, H. J., Goshen-Gottstein, Y., & Usher, M. (2006). Semantic similarity dissociates short- from long-term recency: Testing a neurocomputational model of list memory. *Memory & Cognition, 34,* 323–334.

Elhalal, A. (2008). *Memory processes in list-memory: Experimental and computational investigations.* Unpublished doctoral dissertation, Birkbeck College, University of London, England.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128,* 309–331.

Estes, W. K. (1955). Statistical theory of distributional phenomena in learning. *Psychological Review, 62,* 369–377.

Farrell, S. (2008). *Dissociating conditional recency in immediate and delayed free recall: A challenge for unitary models of recency*. Manuscript submitted for publication.

Farrell, S., & Lewandowsky, S. (in press). Empirical and theoretical limits on lag recency in free recall. *Psychonomic Bulletin & Review*.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 351–360.

Glanzer, M., Gianutsos, R., & Dubin, S. (1969). The removal of items from short-term storage. *Journal of Verbal Learning and Verbal Behavior, 8,* 435–447.

Glenberg, A. M., Bradley, M. M., Kraus, T. A., & Renzaglia, G. J. (1983). Studies of the long-term recency effect: Support for a contextually guided retrieval hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 231–255.

Glenberg, A. M., & Swanson, N. C. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 3–15.

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

Greene, R. L. (1986b). Sources of recency effects in free recall. *Psychological Bulletin, 99,* 221–228.

Haarmann, H. J., Davelaar, E. J., & Usher, M. (2003). Individual differences in semantic short-term memory capacity and reading comprehension. *Journal of Memory and Language, 48,* 320–345.

Haarmann, H. J., & Usher, M. (2001). Maintenance of semantic information in capacity-limited item short-term memory. *Psychonomic Bulletin & Review, 8,* 568–578.

Hawkins, D. A., & Davelaar, E. J. (2005, January). *A closer look at presentation rate effects in free recall.* Poster presented at the meeting of the Experimental Psychological Society, London, England.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory.* New York: Wiley.

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

Howard, M. W., Fotedar, M. S., Datey, A. V., & Hasselmo, M. E. (2005). The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review, 112,* 75–116.

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.

James, W. (1890). *Principles of psychology.* New York: Holt.

Luck, S. J., & Vogel, E. K. (1997, November 20). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–281.

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language, 41,* 3–29.

Martin, R. C., Shelton, J. R., & Yaffee, L. S. (1994). Language processing and working memory: Neuropsychological evidence for separate phonological and semantic capacities. *Journal of Memory and Language, 33,* 83–111.

Mensink, G.-J., & 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 Verbal Learning and Verbal Behavior, 64,* 482–488.

Murdock, B. B., & Okada, R. (1970). Interresponse times in single-trial free recall. *Journal of Experimental Psychology, 86,* 263–267.

Neath, I., & Brown, G. D. A. (2006). SIMPLE: Further applications of a local distinctiveness model of memory. In B. H. Ross (Ed.), *Psychology of learning and motivation* (pp. 201–243). San Diego, CA: Academic Press.

Neath, I., & Crowder, R. G. (1990). Schedules of presentation and temporal distinctiveness in human memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16,* 316–327.

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

Oberauer, K. (2002). Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 411–421.

Polyn, S. M., Norman, K. A., & Kahana, M. J. (2008). *Episodic and semantic organization during free recall: The control of memory search.* Manuscript submitted for publication.

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

Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 1–19.

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

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

Talmi, D., Grady, C. L., Goshen-Gottstein, Y., & Moscovitch, M. (2005). Neuroimaging the serial position curve. A test of single-store versus dual-store models. *Psychological Science, 16,* 716–723.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80,* 352–373.

Usher, M., & Cohen, J. D. (1999). Short-term memory and selection processes in a frontal-lobe model. In D. Heinke, G. W. Humphries, & A. Olsen (Eds.), *Connectionist models in cognitive neuroscience* (pp. 78–91). New York: Springer-Verlag.

Wang, X.-J. (2001). Synaptic reverberation underlying mnemonic persistent activity. *Trends in Neurosciences, 24,* 455–463.

Warden, M. R., & Miller, E. K. (2007). The representation of multiple objects in prefrontal neuronal delay activity. *Cerebral Cortex, 17,* i41–i50.

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

Received March 24, 2008
Revision received May 21, 2008
Accepted May 21, 2008 ■

Postscript: Through TCM, STM Shines Bright

Eddy J. Davelaar and Marius Usher
University of London

Henk J. Haarmann
University of Maryland

Yonatan Goshen-Gottstein
Tel Aviv University

We find the reply by Kahana, Sederberg, and Howard (2008) helpful in clarifying the temporal-context model (TCM) function, in particular with regard to the elimination of the recency effect by a difficult distractor under parameters that still enable long-term contiguity effects to emerge. We agree with Kahana et al. that what matters most to the understanding of memory is the testing of models against actual data, while attempting to maintain the criterion of parsimony. We welcome, therefore, the challenge offered by this exchange, which has produced quite a number of novel predictions (see below). Still, we are not convinced that TCM has been successful in offering a satisfactory account for memory dissociations between long- and short-term recency, that it is able to flexibly discriminate and recall items from different lists, or that it is more parsimonious than is our dual-store model. Our arguments have implications for the wider debate about short-term memory (STM) and long-term memory (LTM).

TCM is unsuccessful in providing an accurate account of the data in several ways. First, out of a number of dissociations between immediate and continuous-distractor free recall (CDR), TCM was able to produce a good account for, at best, one disso-