Reply to Farrell & Lewandowsky: Changes in the shape of the lag-CRP predicted by TCM due to recency

Marc W. Howard
Department of Psychology
Syracuse University
Syracuse, NY 13244, USA

Per B. Sederberg
Department of Psychology and
Center for the Study of Brain, Mind, and Behavior
Princeton University
Princeton, NJ 08540, USA

Michael J. Kahana
Department of Psychology
University of Pennsylvania
Philadelphia, PA 19104, USA

Submitted to *PB&R*, June 25, 2008

Abstract

In the temporal context model (TCM), the current state of context is used as a cue for episodic recall. Farrell & Lewandowsky (in press) argue that the lag-CRP should be examined over a much wider range of lags than have previously been considered. Farrell and Lewandowsky (in press) show that TCM predicts a characteristic change in the shape of the conditional response probability as a function of lag (lag-CRP). This change manifests as a non-monotonicity at extreme lags as well as a skew favoring forward recalls. We show that TCM predicts the distortion to the extent that end-of-list context persists as a retrieval cue for subsequent recall attempts and to the extent that end-of-list context generates a recency effect. Empirically, the degree of skew and non-monotonicity in the lag-CRP seem to be more prominent in immediate than delayed free recall and more prominent in continuous-distractor than delayed free recall. There even appear to be skew and non-monotonicity across lists in a final free recall experiment that exhibited a strong recency effect across lists. TCM predicted the existence of these relatively subtle distortions of the lag-CRP and their correlation with the recency effect. Rather than a reason to suspend work on TCM, these effects provide strong support for an associative engine based on retrieved temporal context (e.g., Sederberg, Howard & Kahana, in press).

In free recall, participants are presented with a list of words and then instructed to recall them in the order they come to mind. Because the order of recall is unconstrained by the experimenter, regularities in the transition probabilities presumably reflect properties of the organization of memory. Perhaps the most important of these regularities in constraining models of episodic memory retrieval is the conditional response probability as a function of lag, or lag-CRP (Kahana, 1996). Given that a participant has just recalled the item from serial position i, the lag-CRP estimates the probability that the next item recalled will be $i + \text{lag}$, attempting to control for the availability of potential recalls in a number of ways. In delayed recall studies, the lag-CRP has a canonical shape, exhibiting a strong contiguity effect favoring adjacent transitions over more remote transitions and an asymmetry favoring forward transitions over remote transitions (see Kahana, Howard, & Polyn, 2008, for a review).

Farrell and Lewandowsky (in press) have identified a novel set of predictions of the temporal context model (TCM, Howard & Kahana, 2002). Previous work has explored the ability of TCM to account for the effects of hippocampal lesion on temporally-defined associations (Howard, Fotedar, Datey, & Hasselmo, 2005), the effect of aging on the shape of lag-CRP curves (Howard, Kahana, & Wingfield, 2006) and the dynamics of immediate and continuous-distractor free recall (Sederberg, Howard, & Kahana, in press). Nonetheless, this subtle prediction about the shape of the lag-CRP curve has not previously been reported, nor have tests of this prediction been undertaken.

Farrell and Lewandowsky (in press) point out that TCM predicts that the persistence of the recency effect across multiple retrieval attempts should lead to a distortion of the lag-CRP. That is, to the extent there is a recency effect, transitions to nearby items should be supplemented with transitions to items at the end of the list. This recency effect leads to a bias towards forward transitions that can manifest as a non-monotonicity in the lag-CRP. If we follow the forward lag-CRP outward from zero, eventually the tendency to make recalls to the end of the list overcomes the advantage from being nearby the just-recalled word. This results in an increase in the lag-CRP at extreme values of lag. The non-monotonicity in the lag-CRP reflects an excess of transitions from extreme serial positions to other extreme serial positions. A persistent primacy effect would manifest as an increase in the lag-CRP in the backward direction—extreme negative lags—whereas a persistent recency effect would manifest as an increase at extreme positive lags.

The persistence of the primacy effect in recall transitions has been known for some time. Laming (unpublished manuscript) observed that there was an excess of transitions to the first serial position—we noted this persistent primacy effect early on in describing the lag-CRP (p. 939, Howard & Kahana, 1999). A moment's reflection reveals that the existence of the primacy effect in the serial position curve obtained in immediate free recall, coupled with the tendency to initiate free recall from the recency portion of the list necessitates an excess of remote transitions to the early part of the list across subsequent retrieval attempts. That is, to the extent that the primacy effect in the serial position curve is not

The authors acknowledge support from National Institutes of Health research grants MH069938 to MWH, MH072138 and MH080526 to PBS and MH055587 MJK. Correspondence concerning this article may be addressed to Marc Howard (`marc@memory.syr.edu`).

solely attributable to the tendency to initiate recall with the first item in the list, then there must have been an excess of remote transitions to the first item.

Farrell & Lewandowsky's primary empirical contribution is to suggest that there is a non-monotonicity in the forward direction consistent with what would be expected from a persistent recency effect. Although the statistics reported by Farrell & Lewandowsky (in press) are not appropriate to determine whether a particular experiment demonstrates a non-monotonicity (see Appendix 1), there are three additional sources of evidence that convince us that lag-CRPs exhibit non-monotonicity at extreme positive values of lag. These sources of evidence are Farrell and Lewandowsky's (in press) meta-analyses (their Figure 2), their observation that the version of TCM they refer to as $\mathrm{TCM_{evo}}$ provides a better fit than the model they refer to as $\mathrm{TCM_{pub}}$ across a wide variety of experiments, and our own secondary analyses (reported here).

In this reply we explore the variables that affect the change of shape of the lag-CRP by examining lag-CRPs from a set of experiments with largely similar methods but differing delay schedules. Qualitative modeling assesses the degree to which this pattern of results is consistent with the predictions of TCM. We start by describing in more detail the source of the distortions in extreme values of the lag-CRP predicted by TCM.

Distortions in the lag-CRP predicted by TCM

TCM proposes that the cue for episodic recall is the current state of a gradually-changing temporal context vector. Potential recalls are cued by a state of context to the extent that it overlaps with the context that obtained when they were studied. The current state of temporal context is driven by presented items, which can also recover their study context. This enables the model to account for contiguity effects—when a studied item is recalled, the input it causes to the temporal context vector resembles the encoding context of neighboring list items, resulting in an increased tendency to recall neighbors of the recalled item. These basic ideas are common to all of the studies that have applied TCM to a variety of topics, although these treatments have varied in a number of details (see Howard & Kahana, 2002; Howard et al., 2005, 2006; Rao & Howard, 2008; Sederberg et al., in press).

TCM predicts a distortion in the shape of the lag-CRP evident in extreme lags to the extent that end-of-list context persists as part of the retrieval cue and to the extent that this end-of-list context supports a recency effect.[1] In TCM, the degree of contextual drift at any given time step is a function of the amount of information that is provided as input to temporal context. This leads to the interesting property that when no input is provided, there is no change in the state of temporal context, predicting that recency can remain intact in response to an unfilled delay (Baddeley & Hitch, 1977; Murdock, 1963). It is perfectly reasonable to suppose that the amount of information provided as

[1]While the primacy effect has been identified with rehearsal (e.g., Brodie & Murdock, 1977; Rundus, 1971; Tan & Ward, 2000), there is, in addition, a one-position primacy effect that is observed in the PFR and remote recall transitions in many data sets (see e.g., Figure 2). While we have occasionally included a descriptive model of primacy in treatments of TCM (e.g., Howard, et al., 2006; Sederberg, et al., in press), this rehearsal-resistant primacy effect is not an integral part of TCM, at least as currently formulated. It is simple enough to add a descriptive account of primacy to TCM to account for primacy in the PFR and persistent backward non-monotonicities in the lag-CRP, such that the existence of primacy does not place a strong constraint on the model. We will not consider primacy further here.

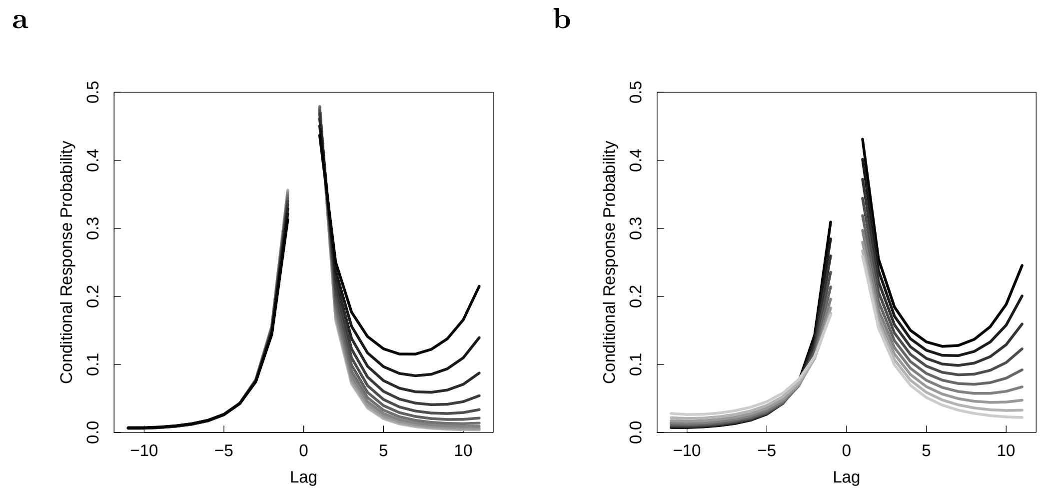

Figure 1. TCM can predict non-monotonicities and skew in the lag-CRP. a. Continuity between the simplified version of TCM (Howard, et al., 2006; Howard, 2004) and the version with retrieval of end-of-list context (Howard & Kahana, 2002). CRP curves were generated at the first output position with $\rho_{\text{ret}} = \rho_{\text{study}}$ (black) and gradually decreasing values of ρ_{ret}, ending with $\rho_{\text{ret}} = 0$, for lighter shaded lines. b. Even without allowing ρ_{ret} to vary from ρ_{study}, TCM can predict a variety of non-monotonicities if the recency effect is altered. The black curves shows the CRP observed at the first output position in immediate free recall with $\rho = .85$, $\gamma = 0.8$, and $\tau = 0.3$. The value of the retention interval was gradually increased from zero to an infinite delay (successively lighter lines).

input following successful retrieval of a memory probe is different than the amount of input caused by study of an item. Thus, the rate of contextual drift during study, ρ_{study}, may differ from the rate of contextual drift during retrieval, ρ_{test}. Farrell and Lewandowsky (in press) consider the case in which $\rho_{\text{test}} = 0$, which they refer to as TCM_{pub}[2] and the case in which $\rho_{\text{test}} = \rho_{\text{study}}$, which they refer to as TCM_{evo}, but in fact there is a continuity of models possible between these when ρ_{test} is allowed to vary independently of ρ_{study} (Polyn, Norman, & Kahana, submitted).

Figure 1a illustrates the predictions of TCM for the shape of the lag-CRP across the entire range of lags for immediate free recall for a variety of values of ρ_{test}. The rate of drift during test controls how much end-of-list context contributes to the retrieval cue at subsequent retrieval attempts. As can be seen from Figure 1a, TCM predicts that to the extent end-of-list context persists during retrieval, the lag-CRP shows an increasingly strong distortion. This is evidenced not only by the non-monotonicity at extreme forward lags, but also by a skew between forward and backward retrievals. Put another way, the model predicts that when end-of-list context does not contribute to retrieval, the asymmetry between forward and backward transitions is approximately constant as the absolute value of lag increases. In contrast, when end-of-list context persists and contributes to the retrieval

[2]Careful examination of Figure 3 of Howard and Kahana (2002) demonstrates a skew and slight non-monotonicity in the lag-CRPs predicted in the original treatment of TCM.

cue, the difference between forward and backward retrievals grows as the absolute value of lag increases. In the backward direction, retrieved context and end-of-list context both favor recall of an item contiguous to a just-recalled item from the middle of the list. In contrast, for retrievals in the forward direction, these cues are in conflict, resulting in a characteristic distortion and even a non-monotonicity.

Even if end-of-list context contributes to subsequent retrievals, TCM predicts forward non-monotonicities only to the extent that end-of-list context is an effective cue for recall of items from the end of the list. Figure 1b shows the lag-CRP predicted by TCM for immediate free recall (black curve) and increasingly long retention intervals (successively lighter curves) when $\rho_{\text{test}} = \rho_{\text{study}}$, i.e., the extreme case referred to as TCM_{evo} by Farrell and Lewandowsky (in press). It is well-known that increasing the retention interval between study of the last item and test results in a decrease of the recency effect in free recall (e.g. Postman & Phillips, 1965). As can be seen from Figure 1b, skew and non-monotonicity are present when the test is immediate and gradually decrease with the increase in the retention interval. Even if retrieval of items during test causes precisely the same amount of input to temporal context as encoding of items during study, the model predicts skew and non-monotonicity only to the extent that the end-of-list context gives rise to a recency effect.

Note that the results of Figure 1b falsify one of the claims of Farrell and Lewandowsky (in press) regarding the predictions of TCM:

> An initial examination of the model revealed a striking non-monotonicity of the forward lag-CRP functions in TCM_{evo}. Irrespective of whether recall was immediate or delayed or involved a continuous distractor task, lags greater than 5 attracted nearly as many—or indeed more—transitions than lags +1.

The light grey lines in Figure 1b, generated with $\rho_{\text{study}} = \rho_{\text{test}}$, which Farrell and Lewandowsky (in press) refer to as TCM_{evo}, show predictions for delayed free recall from TCM. The light grey curves are monotonically decreasing. At extreme lags they do not in any way approach—let alone exceed—the much higher values observed at lags near zero. Farrell and Lewandowsky's (in press) conclusion may be a consequence of fixing the effective delay of the retention interval and/or insufficiently exploring the range of values ρ can take on.

It is important to note that non-monotonicity at extreme lags is not the only issue in differentiating the predictions of TCM when end-of-list context is allowed to persist as a cue from TCM when end-of-list context is not allowed to persist as a cue. This is particularly relevant from an empirical perspective because only the most extreme lags exhibit non-monotonicity (see Figure 2 in Farrell & Lewandowsky, in press). These lags are infrequently observed. For instance, a lag of +11 in a 12-item list, which reflects a transition from the very first item in the list to the very last item in the list—can only be observed if two conditions are met. The very first item in the list must have been recalled and the very last item in the list must be available as a newly-recalled item. In immediate free recall, in which very strong recency effects obtain, these conditions are infrequently met, especially early in output, resulting in a paucity of observations. If limiting one's attention to the first recall transition, these extreme transitions are only observed to the extent that subjects initiate recall with the very first item, which may reflect a serial recall strategy (Bhatarah, Ward, & Tan, 2008).

A qualitative exploration of distortions in the lag-CRP across delay conditions

Appropriate statistical tools have not yet been developed to fully characterize the skew and non-monotonicities in the lag-CRP (see Appendix 1). A qualitative approach to understanding the persistence is more appropriate at this stage of development. Indeed, Farrell and Lewandowsky (in press) use a qualitative approach to good effect in their Figure 2, which calculates an end-adjusted lag-CRP across a broad variety of experiments that vary in delay schedule, list length, modality of presentation, presence of orienting task, presentation rate, level of practice, individual *vs* group testing, and method of recall (verbal *vs* written). Here we will examine the effects of different experimental manipulations on the non-monotonicity, or skew, in the lag-CRP. In order to equate as many variables as possible, we will restrict our attention to published experiments from our labs in which relatively short lists of words were presented visually under conditions designed to minimize rehearsal and verbal free recall was collected. These analyses utilize a subset of the experiments examined by Farrell and Lewandowsky (in press), plus final free recall data (reported in Howard, Youker, & Venkatadass, 2008) from the Howard, Venkatadass, Norman, and Kahana (2007) immediate free recall study. In an attempt to minimize the combinatorial problems associated with missing observations, we will examine the lag-CRPs collapsed across output positions rather than only the first output position. Our analyses suggest that skew and non-monotonicity in the lag-CRP are observed to the extent that there is a recency effect, regardless of whether recall is immediate or delayed. In order to establish this, we first summarize the recency effect observed in these studies.

Recency across delay conditions.

Figure 2 shows the recency effect observed in the experimental data we will consider here. Figure 2a shows the probability of first recall (PFR) curves from Experiments 1 and 2 of Howard and Kahana (1999), exhibiting immediate, delayed and continuous-distractor free recall. In immediate free recall, the list is presented and test immediately follows presentation of the last item. A strong recency effect is observed in the PFR. In delayed free recall, a delay intervenes between study of the last item and the test. In delayed free recall, the recency effect is attenuated relative to immediate free recall. In continuous-distractor free recall, a delay intervenes between each list item and also at the end of the list. The recency effect in the PFR is larger in continuous-distractor free recall than in delayed free recall. The long-term recency effect in continuous-distractor free recall is also observed when examining final free recall across lists (Tzeng, 1973; Glenberg et al., 1980). Howard et al. (2007) examined immediate free recall of 48 lists. A recency was observed in immediate free recall testing. At the end of the session, they tested final free recall of all the items from all the lists. Howard et al. (2008) reported the final free recall results from this study and observed a recency effect across lists in the PFR (Figure 2b). Notably, there was no recency effect relative to within-list serial position in final free recall. This makes sense in that the delay between study of a particular list and the final free recall period could be several tens of minutes.

The lag-CRP at extreme values across delay conditions.

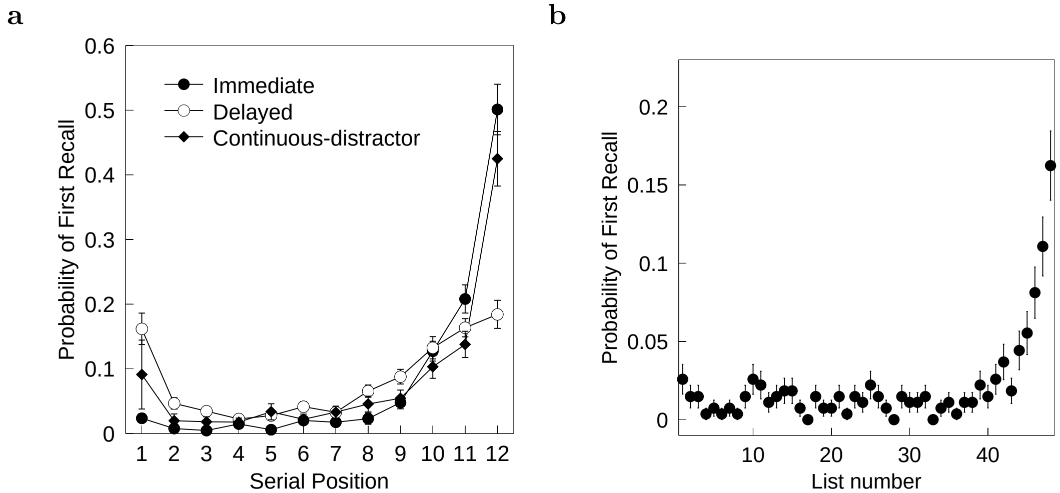

Figure 2. The recency effect across delay schedules. a. The recency effect, as illustrated by the probability of first recall (PFR), from immediate, delayed and continuous-distractor free recall from Experiments 1 and 2 from Howard & Kahana (1999). While the recency effect is attenuated in delayed free recall, it is amplified in continuous distractor free recall. b. The recency effect in the PFR across lists from Howard, Youker, & Venkatadass (2008). In both panels, error bars reflect the standard error of the mean.

Figure 3 illustrates lag-CRPs from immediate, delayed and continuous-distractor free recall, as well as final free recall across lists. While prior work has focused on documenting the existence of a contiguity effect by focusing on lags around zero (e.g. Howard & Kahana, 1999; Howard et al., 2008), here we examine all possible lags, as suggested by Farrell and Lewandowsky (in press). Figure 3a compares immediate to delayed free recall from Experiment 1 of Howard and Kahana (1999). There is a boost in the contiguity effect and appears to be a larger non-monotonicity in immediate free recall compared to delayed free recall.

Figure 3b compares the longest-IPI condition of Experiment 2 of Howard and Kahana (1999), labeled “continuous-distractor,” to the zero-IPI condition, labeled “delayed.” Recall that continuous-distractor free recall shows a larger recency effect than delayed free recall (Figure 2a). Although the contiguity effect is similar in magnitude across the conditions, the non-monotonicity exhibited at extreme positive lags appears stronger in continuous-distractor free recall than in delayed free recall.

Figure 3c compares the lag-CRP from immediate free recall in the Howard et al. (2007) data with the within-list lag-CRP observed in final free recall of the same items. The final free recall data is comparable to delayed free recall with an extremely long delay. The results appear to be consistent with Figure 3a. Again the contiguity effect is larger in magnitude in immediate recall. In addition a non-monotonicity in the forward direction is observed in immediate free recall. With the delay, leaving aside the primacy effect observed at extreme backward lags, the lag-CRP has a consistent degree of asymmetry in contrast to the skew observed in the immediate free recall data. Figure 3d illustrates the *across-list* lag-CRP observed by Howard et al. (2008). There appears to be a steep non-monotonicity

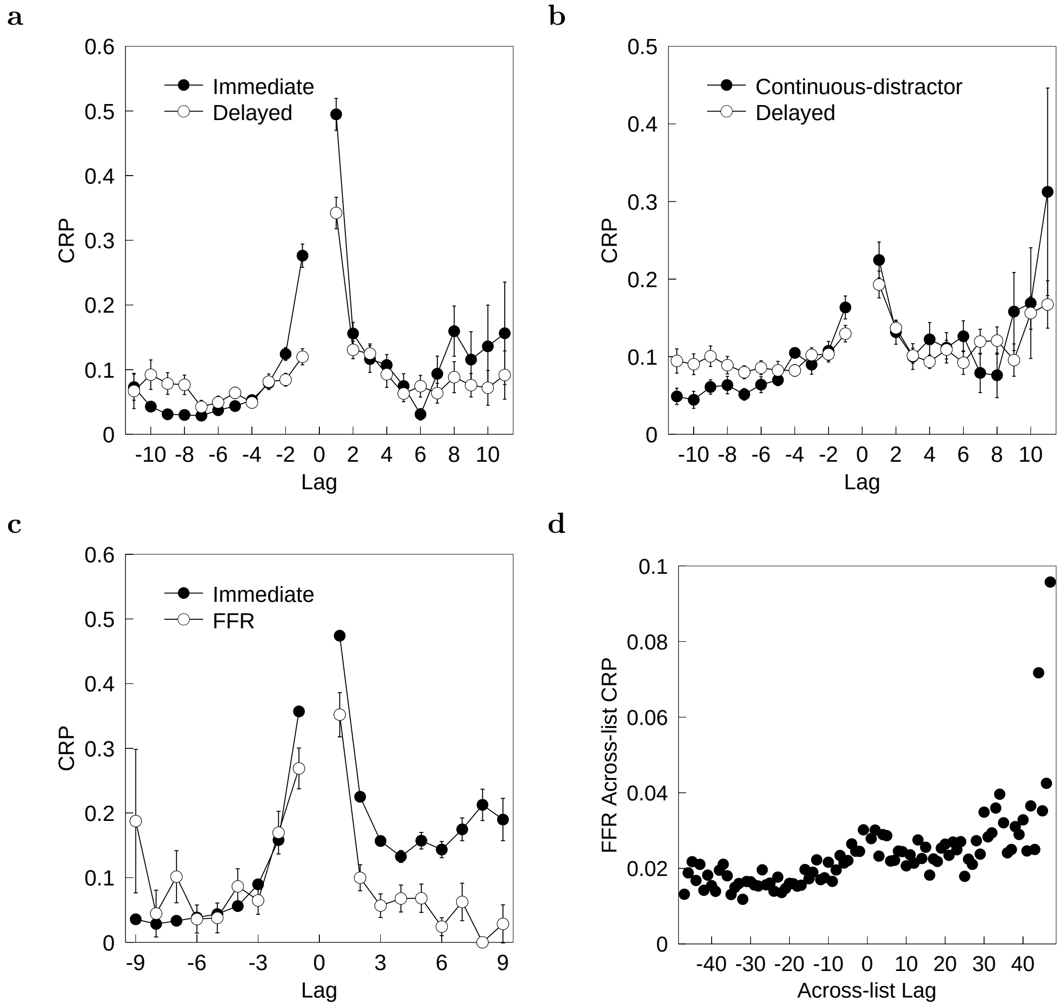

Figure 3. Non-montonicities in the lag-CRP appears to co-occur with the recency effect. In all panels, lag-CRPs are calculated across all output positions. a. Immediate condition and delayed condition of Experiment 1 of Howard & Kahana (1999). b. Comparison of longest-IPI condition (labeled "Continuous-distractor") and zero-IPI condition ("Delayed") of Experiment 2 of Howard & Kahana (1999). c. Lag-CRP from the immediate free recall test of the control lists of Howard, Venkatadass, Norman & Kahana (2007) contrasted with the within-list CRP calculated from a final free recall session ("FFR"). Error bars in panels a-c reflect one standard error. d. Across-list CRP (Howard, Youker, & Venkatadass, 2008).

observed at extreme positive lags presumably corresponding to the across-list recency effect observed in the same data (Figure 2b).

Qualitative predictions of TCM

Here we consider whether the properties of the the lag-CRP across the entire range of possible lags are consistent with the predictions of TCM. Our strategy, is to use a common set of parameters that illustrate the qualitative behavior of the model across conditions, which we will compare to the pattern of observed results across experiments. The goal of this approach is to provide insight as to whether the source of the skew and non-monotonicity in the lag-CRP are consistent with the origin predicted by TCM.

Farrell and Lewandowsky (in press) evaluated fits of a two-parameter TCM model and found that the predictions of the model deviated from the observed results to an extent significantly different from chance. This is not a surprising result; there are many sources of variability that are not included in this two-parameter description. For instance, it is known that free recall is strongly affected by the degree of proactive interference the items are subject to (Goodwin, 1976), the duration of the delay interval (Postman & Phillips, 1965), and the semantic organization of the list (Romney, Brewer, & Batchelder, 1993; Glanzer, Koppenaal, & Nelson, 1972). The failure of a two-parameter model to account for all of these sources of variability is not at all troubling.

Indeed, free recall models that have been evaluated in recent years have typically used a great many more than two parameters and been content to describe the qualitative pattern of results across experiments. Following early work on the SAM model (Raaijmakers & Shiffrin, 1980), these models have typically used a common set of parameters that provide a qualitative description of a broad range of phenomena. For instance, the Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, and Usher (2005) paper described a model that had thirteen free parameters and provided no formal goodness-of-fit statistics. Similarly, the TCM-A model of Sederberg et al. (in press) has eight free parameters, four fixed parameters, and reports no goodness-of-fit statistics. Recent variants of SAM applied to free recall (Sirotin, Kimball, & Kahana, 2005; Kimball, Smith, & Kahana, 2007) have ten or more free parameters. While the fSAM model of Kimball et al. (2007) made extensive use of goodness-of-fit statistics to compare model variants to one another, there was no attempt in either paper to compare the model to the data in an absolute sense.[3] Memory researchers have long appreciated the richness and complexity of the free recall task. Accordingly, models of free recall have placed a premium on trying to gain insight into the basic mechanisms that underlie memory retrieval rather than curve-fitting.

In evaluating the qualitative predictions of TCM, we informally searched for a set of parameters that would exhibit the basic properties of the recency effect and lag-CRP effects exhibited by the data across conditions from the single-list free recall experiments (Figures 2a and 3a-c). A relatively broad range of parameters exhibit the same basic properties. These parameters, with $\rho = .85$, $\gamma = .8$, $\tau = .3$ and $\rho_D = 0.4$ were used in

[3] Brown, Neath and Chater's (2007) SIMPLE model is something of an exception to this pattern, in that it has only three free parameters that were varied across experiments. SIMPLE is consistent with the pattern in that it was evaluated informally using R^2. It should be noted that while SIMPLE has been applied to serial position curves in free recall, as well as data from a wide variety of other memory and discrimination tasks, it has not been applied to CRP curves or other aspects of the dynamics of free recall.

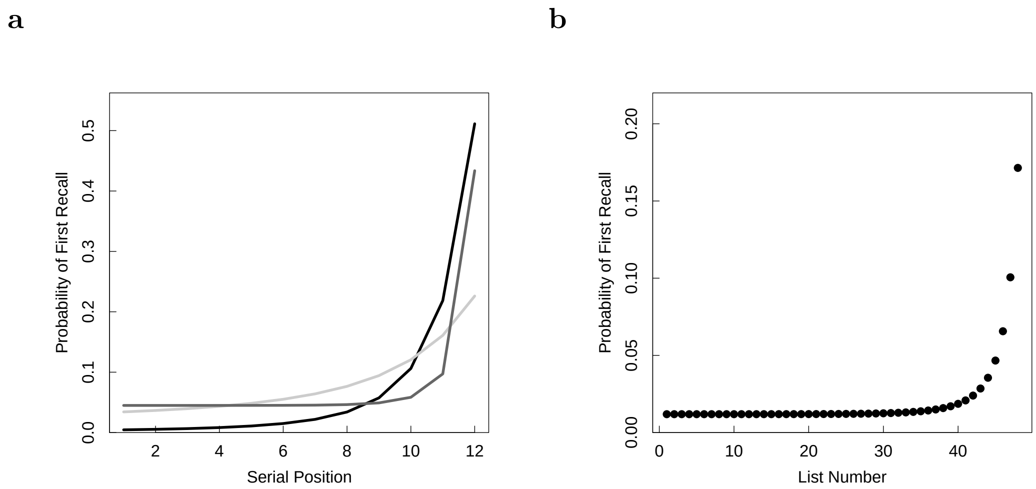

Figure 4. TCM with the Luce choice retrieval rule produces recency effects that correspond to the qualitative pattern observed. Compare to Figure 2. a. Probability of first recall functions for immediate (black), delayed (light grey) and continuous-distractor (dark grey) free recall. The same parameters were used in Figure 5a-c. b. Across-list PFR. The same parameters were used in Figure 5d.

the predictions generated in Figures 1a and b, Figure 2a and Figures 3a-c (see Appendix 2 for details of the parameterization). Because we treated the lists as single items in the across-list FFR simulations, a separate set of parameters were chosen for the across-list free recall data, with the effective rate of contextual drift across lists set to .8, $\gamma = .8$, τ during the PFR set to .6 and τ for the CRP set to .9.[4] In all of these predictions, ρ_{study} was equal to ρ_{test}, consistent with the formulation Farrell and Lewandowsky (in press) referred to as $\mathrm{TCM}_{\mathrm{evo}}$.

Figure 4a shows the predictions of TCM for the PFR from immediate, delayed and continuous-distractor free recall. As can be seen by comparing Figure 4a to the empirical results shown in Figure 2a, the model correctly predicts a strong recency effect in immediate free recall, an attenuation of the recency effect in delayed free recall and a strong recency effect in continuous-distractor free recall (see also Howard & Kahana, 2002). In modeling the recency effect in final free recall across lists, TCM can also predict a recency effect that extends across multiple lists.

To avoid issues with modeling resampling, latency and recall cessation, we generated predictions for the lag-CRP at the first output position from TCM. Some caution should be

[4]The point of fitting the across-list PFR and lag-CRP here is simply to illustrate that TCM predicts that the non-monotonicity in the lag-CRP should be correlated with the recency effect in the PFR, which appears to be supported by the empirical findings. The change in τ across retrieval attempts can be justified as a result of items retrieving noise during successive retrieval attempts. A more accurate treatment of this experiment would take into account that the delay between study of the last list and the FFR session was longer than the delay between lists, the effect of immediate recall as an encoding event, and the fact that with a total of several hundred items presented across 48 lists, the assumption that all item vectors are orthogonal becomes increasingly untenable.

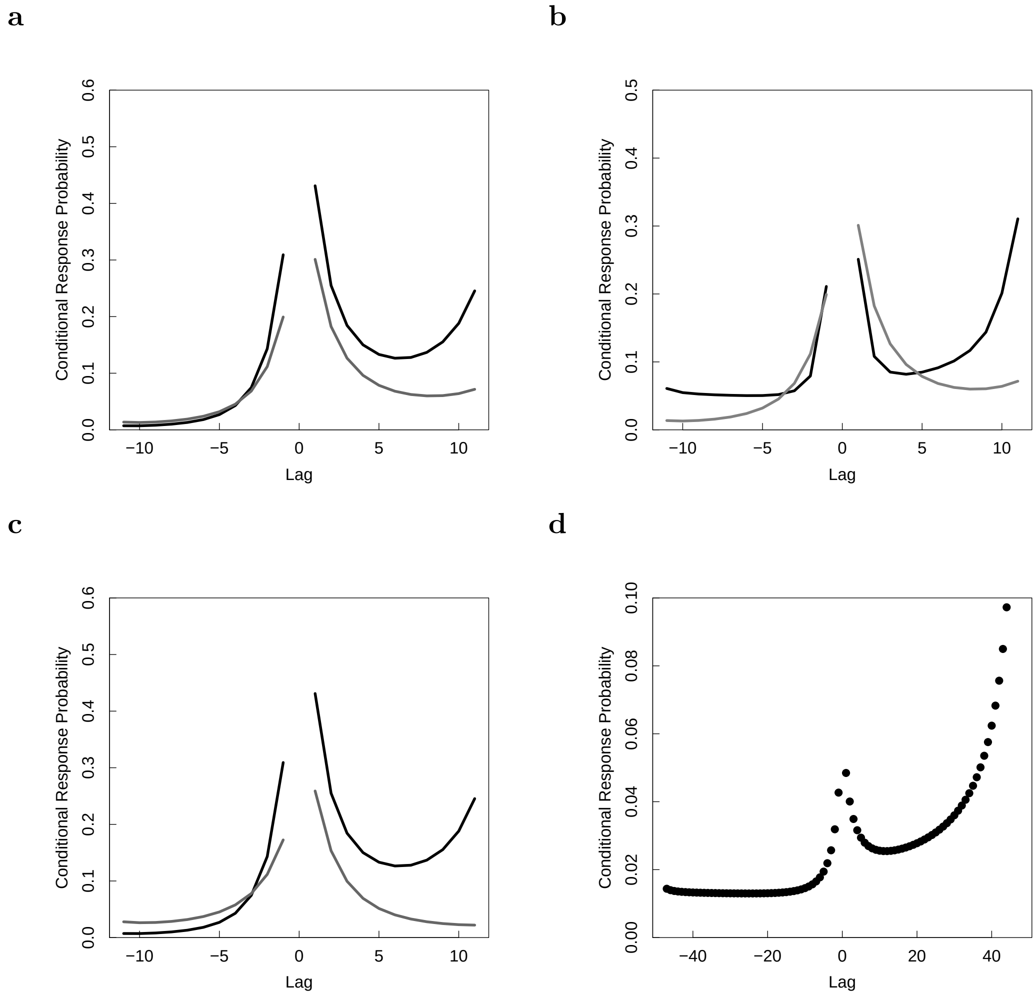

Figure 5. Simulations of lag-CRPs from the first output position using TCM, with the Luce choice retrieval rule. Compare to Figure 3. a. Immediate free recall (black) and delayed free recall. b. Continuous-distractor free recall (black) and delayed free recall (grey). c. Immediate free recall compared to delayed free recall with an infinite delay. d. Across-list CRP.

exercised in comparing the predictions of the model with the empirical results in Figure 3, which collapse across output positions. In order to successfully model lag-CRPs across more than the first output transition, one must have a mechanism for dealing with resampling and termination of recall, neither of which depend on the structural properties of TCM. While the comparison of the qualitative shape of the lag-CRP is not dramatically altered by collapsing across output position, the magnitude of the contiguity effect in immediate free recall is affected by output position (Howard & Kahana, 1999; Kahana, Howard, Zaromb, & Wingfield, 2002). As has been well-known for some time, TCM with the Luce choice rule under-predicts the magnitude of the lag-CRP effect at early output positions in immediate free recall (Howard & Kahana, 2002). The use of lag-CRPs from only one output position in the modeling and the use of lag-CRPs collapsed across output positions tends to obscure this difference in immediate free recall.

Figure 5 shows the lag-CRPs from the first output position from the same experimental settings that generated the corresponding PFR curves in Figure 4. As can be seen from Figure 5a, TCM predicts that the lag-CRP from immediate free recall should have a larger contiguity effect, somewhat greater skew and a larger non-monotonicity in the forward direction than the lag-CRP from delayed free recall. In continuous-distractor free recall (Figure 5b), the model correctly predicts that although the contiguity effect should be of similar magnitude across conditions, the difference between the lag-CRP from continuous-distractor free recall and delayed free recall should be manifest as a large and sharp non-monotonicity at extreme positive lags in continuous-distractor free recall relative to delayed free recall.[5]

Figure 5c shows predictions comparing the lag-CRP from immediate free recall within-list lag-CRP from final free recall of the same lists. To simulate the very long delay between study of a typical list from the experiment and the final free recall session, we simply set the effective length of the retention interval to be infinite in generating the grey lag-CRP curve in Figure 5c, rather than modeling delayed free recall as reflecting a small residual recency effect (Figures 2a, 4a). Indeed, there is no evidence for a within-list recency effect or primacy effect in these final free recall data (Howard et al., 2008). The correspondence between the predictions (Figure 5c) and the empirical observations (Figure 3c) in this case are particularly strong. The model not only correctly predicts a non-monotonicity in the immediate lag-CRP that is larger than in final free recall, but also describes a continuously-increasing difference between immediate free recall and delayed final free recall for increasing values of lag. Moreover, the model also correctly predicts a benefit for backward transitions in delayed final free recall. It is perhaps worth noting that this was a very large study, with almost three hundred participants performing more than 7,000 trials of immediate free recall.

Comparing Figure 5d, which shows predicted values of across-list lag-CRPs, with Figure 3d, which shows empirically-observed across-list lag-CRPs, we see that the model has successfully captured several aspects of the data. First, the model correctly describes a contiguity effect across several lists that is also exhibited in the data[6] It is worth noting

[5] The TCM-A model of Sederberg, et al. (2008) shows the same pattern of results at extreme lags across delay conditions using the published parameter settings.

[6] Howard, et al. (2008) conducted analyses on a surrogate data set to confirm that the boost in the lag-CRP was not an artifact of a persistent recency effect or in fact any other variation in encoding across

that these large-scale contiguity effects (see also Howard & Kahana, 1999, Figure 3c) are a natural prediction of TCM and are also a challenge for buffer accounts of contiguity effects (see Howard & Kahana, 2002; Davelaar et al., 2005; Sederberg et al., in press). Second, there is a large non-monotonicity in the forward direction, such that extremely large across-list lags are better recalled than adjacent across-list lags. Third, there is a persistent skew to the entire curve that appears to be reflected in the data.

From these analyses, the model appears to correctly predict the range of shapes of lag-CRP curves that are observed, and the variation in the skew and non-monotonicity observed across experiments. In particular, these distortions seem to co-occur with the recency effect, as predicted by TCM. Rather than ruling out TCM as a description of recency and contiguity across scales, the data from examining the entire range of lag-CRPs across conditions are qualitatively quite consistent with the predictions of TCM.

Contiguity effects predicted by TCM in immediate free recall are not an averaging artifact

Farrell and Lewandowsky (in press) claimed that TCM incorrectly predicts an artifactual lag-CRP due to averaging across serial positions with a strong recency effect. In their Figure 7, Farrell and Lewandowsky (in press) examined the lag-CRP conditionalized on the serial position of the previously-recalled word. Although noisy, the experimental data from the Howard et al. (2007) study show a contiguity effect in the forward direction for each previously-recalled serial position with a non-monotonicity at extreme positive lags for some items. In contrast, the best-fitting parameter values of their two-parameter implementation of TCM showed a pure recency effect in the lag-CRP for the forward direction when conditionalized on serial position of the just-recalled item. If this were a general property of the model, it would clearly rule out TCM as a description of the contiguity effect. However, Farrell & Lewandowsky's (in press) finding is an artifact of the particular choice of parameters used in generating the predictions and does not reflect a general property of the model.

Figures 6a and b show the predictions from TCM with the Luce choice retrieval rule using the same parameters as Figures 2a and 3a-c. The model clearly shows a contiguity effect at each serial position in addition to a recency effect that appears as a non-monotonicity at extreme lags.

Although TCM with the Luce choice rule (Howard & Kahana, 2002) does a good job of accounting for the basic pattern of results seen in the PFR and lag-CRPs, it was never intended as a serious model of the first several transitions in the early stages of immediate free recall. Howard and Kahana (2002) specifically argued that it did not properly capture the dramatic boost in the lag-CRP at early stages of immediate free recall (section 5.3, "TCM is not a free recall model," Howard & Kahana, 2002). Important distinctions between the properties of immediate recency and long-term recency were highlighted by Davelaar et al. (2005) in their buffer model of immediate recency and variable context model of long-term recency. Sederberg et al. (in press) showed that these dissociations can be addressed in the framework of TCM if the Luce choice rule were replaced with a set of competing accumulators. Sederberg et al. (in press), who referred to this model as TCM-A, showed that these elaborations of the retrieval rule enabled TCM-A to account for numerous dissociations be-

the experimental session.

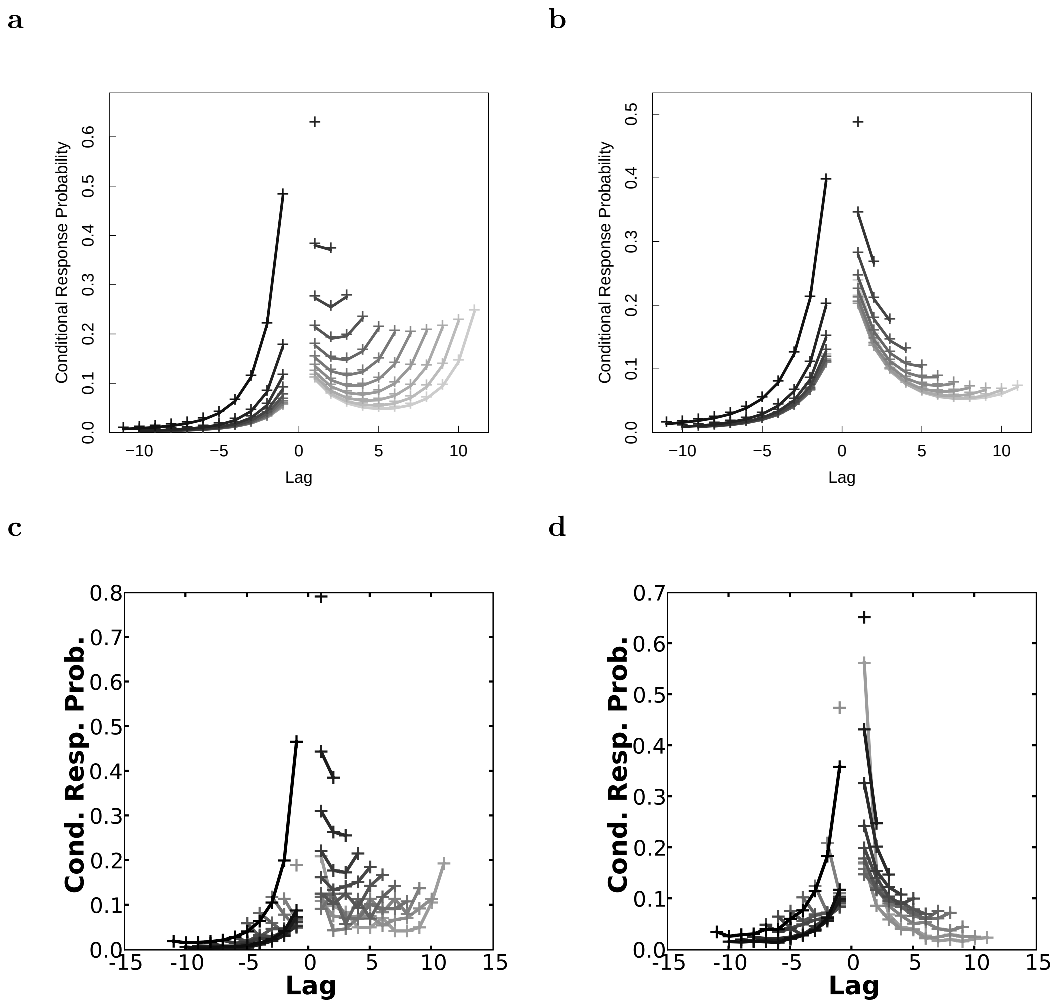

Figure 6. The CRP predicted by TCM in immediate free recall need not be an artifact of the recency effect. Compare to Figure 7, Farrell & Lewandowsky (in press). Panels on the left (a and c) correspond to immediate free recall. Panels on the right (b and d) correspond to delayed free recall (note change of scale). a,b. Simulations from the Howard & Kahana (2002) version of TCM with the Luce choice retrieval rule. The same parameters were used as in Figure 4a and Figure 5a-c. c,d. Simulations from the TCM-A model of Sederberg, Howard & Kahana (in press). CRP curves broken down by serial position of the just-recalled item are shown. Parameter values are as reported in Sederberg et al. (in press).

tween immediate recency and long-term recency, including differences in latency, sensitivity to proactive interference, the differential effects of amnesia on immediate and delayed recall, and differential sensitivity of the contiguity effect to output position across conditions. Interestingly, Sederberg et al. (in press) adopted the convention that $\rho_{\text{study}} = \rho_{\text{test}}$, referred to as TCM_{evo} by Farrell and Lewandowsky (in press).

TCM-A (Sederberg et al., in press) constitutes a serious attempt to describe the early stages of immediate free recall using retrieved temporal context as the retrieval cue. If TCM-A also leads to an artifactual account of the contiguity effect, like that illustrated by Farrell and Lewandowsky (in press) for their two-parameter implementation of TCM, this would be a serious challenge to the retrieved context framework. To evaluate if this is the case, we ran a simulation of TCM-A with 50,000 simulated trials using the same set of parameters used by Sederberg et al. (in press). Figures 6c and d show the lag-CRP segregated by serial position of the just-recalled word in immediate and delayed free recall respectively. As can be clearly seen from Figure 6c, TCM-A predicts a contiguity effect, as well as a non-monotonicity in the forward direction, even when the lag-CRP is conditionalized on the serial position of the just-recalled item. It should be noted that the parameter settings obtained by Sederberg et al. (in press) were obtained to maximize the model's description of a narrow range of lags across delay conditions. Nonetheless, the model makes accurate qualitative predictions for extreme lags. We conclude that although it is possible to find parameters for TCM with the Luce choice rule—and probably TCM-A as well—that generate an artifactual contiguity effect, this is not a weakness of the model *per se* so much as a weakness of the specific choice of parameters used by Farrell and Lewandowsky (in press).

General Discussion

Farrell and Lewandowsky (in press) pointed out that TCM makes a prediction about the shape of the lag-CRP in free recall when the lag-CRP is considered across all possible lags—even those with very few observations. Their meta-analysis that aggregates lag-CRP curves collected under a wide variety of experimental conditions (Figure 2 of Farrell & Lewandowsky, in press), their finding that the variant of TCM they refer to as TCM_{evo} provided a superior fit than the variant of TCM they refer to as TCM_{pub}, and our own analyses (Figure 3) all suggest that these changes in the shape of the lag-CRP at extreme values of lag are observed—at least under some circumstances. Our own secondary analyses (Figure 3) suggest that the non-monotonicity in the extreme values of the lag-CRP and skew are driven by persistent serial position effects. We focused our attention on the non-monotonicity in the forward direction and the persistent skew in the lag-CRP which are a consequence of a persistent recency effect. TCM successfully describes the conditions under which the recency effect is observed (Figure 4) and, as a consequence, also the qualitative pattern of lag-CRPs observed across conditions (Figure 5).

Farrell & Lewandowsky's (in press) conclusion, "that TCM must await additional development before it can advance our understanding of free recall processes," is, ironically, falsified by their own findings. They describe a prediction of TCM and then demonstrate that it is observed in the data, thus advancing our understanding of free recall. Although we would be the first to agree that TCM is not a complete description of free recall (and nor is TCM-A Sederberg et al., in press, although it is certainly much closer) the ideas at the core of TCM—that a gradually-changing state of temporal context is the cue for episodic recall and that some items can recover the state of temporal context in which they were encoded—have yielded a number of predictions that have advanced our understanding of free recall. In addition to the predictions examined here, a number of other predictions about free recall performance, and episodic memory more broadly, have been confirmed in recent years.

For instance, TCM describes contextual drift during study as a consequence of contextual retrieval. This predicts that repetition of items from the list will produce a transient associative advantage for neighbors of the initial presentation of the repeated item. Howard et al. (2007) observed such associations in the early stages of immediate free recall, including immediate free recall initiation, in the absence of explicit recall of the repeated item. This prediction strongly differentiates TCM from buffer accounts of immediate recency (Atkinson & Shiffrin, 1968; Raaijmakers & Shiffrin, 1980; Davelaar et al., 2005), advancing our understanding of free recall. Other predictions of TCM in paired-associate learning (Howard, Jing, Rao, Provyn, & Datey, revised; Provyn, Sliwinski, & Howard, 2007) and item recognition (Schwartz, Howard, Jing, & Kahana, 2005) have also been experimentally observed, advancing our understanding of episodic associations more broadly.

Toward a free recall model based on TCM

Two detailed free recall models based on TCM have been recently developed. The TCM-A model of Sederberg et al. (in press) provides a description of dissociations between immediate recency and long-term recency. The CMR model of Polyn et al. (submitted) uses a framework very similar to TCM to account for associative effects observed in free recall of words encoded in different task contexts. Farrell and Lewandowsky (in press) have provided a service by pointing out that the entirety of the lag-CRP curve can yield important constraints on models of free recall. However, the constraints appear on further examination to be quite consistent with the qualitative predictions of the model. This does not mean, however, that all of the challenges necessary to build a complete model of free recall based on TCM have been addressed. We note a couple of such challenges here.

A potentially major weakness of the current TCM framework in describing free recall is that it does not provide a principled explanation for the existence of the processes by which participants undertake a strategic memory search. For instance, Bhatarah et al. (2008) presented participants with a brief list and post-cued them for immediate free recall or immediate serial recall. They found that the PFR curves for free recall showed a strong recency effect whereas the PFR curves for serial recall showed a dramatic primacy effect. While it is straightforward to build in extra encoding strength for the first item in the list in TCM, the ability to strategically alter the memory search sufficiently to change the shape of the serial position curve is not part of any of the current implementations of TCM. It may be possible to account for this strategic switch if one supposes that participants somehow insert something analogous to a retention interval before attempting serial recall, or perhaps by adopting a generate/recognize strategy in which they recall several items until they find the first item in the list. Both of these accounts are consistent with the finding that participants are much slower to initiate serial recall than they are to move from item to item (e.g. Farrell & Lewandowsky, 2004; Kahana & Jacobs, 2000). Interestingly, Bhatarah et al. (2008) found lag-CRP curves of the same basic form, with forward and backward contiguity effects and asymmetry effects, across tasks and cueing condition, suggesting that the mechanisms underlying temporally-defined associations are not strongly affected by recall strategy.

References

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, p. 89-105). New York: Academic Press.

Baddeley, A. D., & Hitch, G. J. (1977). Recency reexamined. In S. Dornic (Ed.), *Attention and performance VI* (p. 647-667). Hillsdale, NJ: Erlbaum.

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.

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.

Brown, G. D., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539-76.

Davelaar, E. J., Goshen-Gottstein, Y., Ashkenazi, A., Haarmann, H. J., & Usher, M. (2005). The demise of short-term memory revisited: empirical and computational investigations of recency effects. *Psychological Review*, *112*(1), 3-42.

Farrell, S., & Lewandowsky, S. (2004). Modelling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory and Language*, *51*, 115-135.

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

Glanzer, M., Koppenaal, L., & Nelson, R. (1972). Effects of relations between words on short-term storage and long-term storage. *Journal of Verbal Learning and Verbal Behavior*, *11*, 403-416.

Glenberg, A. M., Bradley, M. M., Stevenson, J. A., Kraus, T. A., Tkachuk, M. J., & Gretz, A. L. (1980). A two-process account of long-term serial position effects. *Journal of Experimental Psychology: Human Learning and Memory*, *6*, 355-369.

Goodwin, C. J. (1976). Changes in primacy and recency with practice in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior*, *15*, 119-132.

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*(1), 75-116.

Howard, M. W., Jing, B., Rao, V. A., Provyn, J. P., & Datey, A. V. (revised). Bridging the gap: Transitive associations between items presented in similar temporal contexts. *Journal of Experimental Psychology : Learning, Memory, and Cognition.*

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*(3), 269-299.

Howard, M. W., Kahana, M. J., & Wingfield, A. (2006). Aging and contextual binding: Modeling recency and lag-recency effects with the temporal context model. *Psychonomic Bulletin & Review*, *13*, 439-445.

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

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

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

Kahana, M. J., Howard, M., & Polyn, S. (2008). Associative processes in episodic memory. In H. L. Roediger III (Ed.), *Cognitive psychology of memory, Vol. 2 of learning and memory - a comprehensive reference (J. Byrne, Editor)* (p. 476-490). 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.

Kahana, M. J., & Jacobs, J. (2000). Inter-response times in serial recall: Effects of intraserial repetition. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, *26*, 1188-1197.

Kimball, D. R., Smith, T. A., & Kahana, M. J. (2007). The fSAM model of false recall. *Psychological Review*, *114*(4), 954-93.

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

Murdock, B. B. (1963). Short-term memory and paired-associate learning. *Journal of Verbal Learning and Verbal Behavior*, *2*, 320-328.

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

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

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

Provyn, J. P., Sliwinski, M. J., & Howard, M. W. (2007). Effects of age on contextually mediated associations in paired associate learning. *Psychology and Aging*, *22*, 846-857.

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.

Rao, V., & Howard, M. (2008). Retrieved context and the discovery of semantic structure. In J. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Advances in neural information processing systems 20.* Cambridge, MA: MIT Press.

Romney, A. K., Brewer, D. D., & Batchelder, W. H. (1993). Predicting clustering from semantic structure. *Psychological Science*, *4*, 28-34.

Rundus, D. (1971). An analysis of rehearsal processes in free recall. *Journal of Experimental Psychology*, *89*, 63-77.

Schwartz, G., Howard, M. W., Jing, B., & Kahana, M. J. (2005). Shadows of the past: Temporal retrieval effects in recognition memory. *Psychological Science*, *16*(11), 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*.

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*, 787-805.

Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology : Learning, Memory, and Cognition*, *26*(6), 1589-1626.

Tzeng, O. J. L. (1973). Positive recency in delayed free recall. *Journal of Verbal Learning and Verbal Behavior*, *12*, 436-439.

Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review*, *11*(1), 192-6.

Appendix 1: Empirical challenges in measuring persistent marginal serial position effects

While we salute Farrell & Lewandowsky's pioneering effort to quantitatively characterize the non-monotonicity in extreme values of the lag-CRP, there are several challenges inherent in these analyses that have not yet been adequately met.

Flaws and limitations in the Farrell & Lewandowsky (in press) analyses.

Farrell and Lewandowsky (in press) quantitatively characterized the degree of non-monotonicity in the lag-CRP observed by fitting various descriptive models to the observed lag-CRPs using maximum likelihood estimation. Two of these models, the quadratic and complementary exponential are capable of exhibiting non-monotonicity and two, the linear and power function, are not. The results of these analyses are summarized in Table 2 of Farrell and Lewandowsky (in press), which provides Akaike weights (Wagenmakers & Farrell, 2004) for each of the models. The Akaike weight can be interpreted as the conditional probability that a particular model is correct, given the data *and the set of models entered into the analysis.*

The first hint that something is wrong with the Farrell and Lewandowsky (in press) empirical analyses comes from the inconsistency, with extremely high confidence, in the verdict of the analyses. For instance, focusing on the forward direction, we find that the probability of one of the non-monotonic models describing the data for the Howard et al. (2007) data is certainty, but the probability of one of the monotonic models (in particular the power function) is also certainty for the Murdock and Okada (1970) data. While there are certainly methodological differences between the two experiments, they are both large studies of immediate free recall, so it seems odd that the results would be so completely discrepant, and with such a high degree of certainty. Even more puzzling are the results for the subconditions of the Murdock (1962) data. If one takes the Akaike weights seriously, they suggest that we can be certain that the 20-item lists presented once per second show a non-monotonicity, as do the 30-item lists, but the 40-item lists just as certainly do not show a non-monotonicity, as do the 20-item lists presented once per two seconds! What would account for such a large discrepancy in the qualitative properties of performance in response to parametric manipulations in the methods of the experiment?

	ΔAIC			
	First output position		Collapsed	
Experiment	Exp.-to-asymp.	Comp. exp.	Exp.-to-asymp.	Comp. exp.
HK99 Exp 1 Immed	145.6	0	0	47.6
HK99 Exp 1 Delay	138.4	0	0	77.4
HK99 Exp 2 IPI=0	0	5.4	0	9.7
HK99 Exp 2 IPI=2	0	20.4	0	18.2
HK99 Exp 2 IPI=8	45.8	0	0	15.4
HK99 Exp 2 IPI=16	167.6	0	0	18.4
HVNK07	0	429	0	308.4

Table 1: Comparison of the complementary exponential model (Comp. exp.) reported by Farrell and Lewandowsky (in press) with an exponential decay to an asymptote (Exp.-to-asymp.). HK99=Howard and Kahana (1999). HVNK07=Howard, Venkatadass, Norman and Kahana, 2007. IPI=Interpresentation interval. ΔAIC=difference between the AIC for the best-fitting model and the model in question.

Upon further reflection, one reason for the wild fluctuation in the Akaike weights is that the set of models under consideration is not appropriate to the empirical question being asked. Farrell and Lewandowsky (in press) evaluated non-monotonicity by comparing three-parameter models that can exhibit non-monotonicity with one parameter models that cannot. However, the ability to exhibit non-monotonicity is not the only way in which these classes of models differ. Most notably, the three-parameter models are able to exhibit a non-zero asymptotic value whereas the one-parameter models cannot. Consider Farrell & Lewandowsky's Eq. 4:

$$P(l) = c \exp(-al) + (1-c) \exp(-bl).$$

This complementary exponential function is nested with a two-parameter model that is not monotonic yet gives rise to a non-zero asymptote that is achieved by setting $b = 0$. This two-parameter model describes an exponential decay to an asymptote. It does not have the ability to increase at long lags, yet provides a degree of flexibility that the "monotonic" functions cannot exhibit.

The analyses used by Farrell and Lewandowsky (in press) are inappropriate to evaluate non-monotonicity. It is possible that the non-monotonic exponential-to-asymptote model provides a better fit than the complementary exponential function. To evaluate this possibility, we compared the exponential-to-asymptote model to the complementary exponential model using the methods described by Farrell and Lewandowsky (in press) in their supplementary material. We examined forward lags only, comparing the models both for the first output position lag-CRP and the lag-CRP collapsed across output positions. For experimental data, we used the Howard et al. (2007) data from the control lists and both experiments of Howard and Kahana (1999), for a total of seven conditions.

Table 1 shows the difference in AIC for the models. The exponential-to-asymptote outperforms the complementary exponential model for 3/7 data sets at the first output position and 7/7 data sets when collapsed across output positions. A simple-minded application of this result may lead us to conclude that the lag-CRP is not non-monotonic. This would be an inappropriate conclusion for several reasons. First, these two descriptive models may be of inappropriate form—perhaps one would find non-monotonicity if one used a complementary power function or some other form entirely. It is also possible that the apparent superiority of the simpler model is an artifact of the different numbers of observations at different lags. For instance, of the 294 participants in Howard et al. (2007), 118 have precisely one non-zero probability over the range of lags from +2 to +9 (these were almost always at lag +2). Because of the way that the log-likelihood is calculated, this means that the function that can produce the highest probability for the lag with non-zero number of observations will produce the best fit for that subject. The predictions of the complementary exponentials model are most distinct from the exponential-to-asymptote model at extreme lags. If those lags are not observed, then the simpler model will be favored.

The weakness of the evidence for non-monotonicity from the complementary exponentials model is not strong evidence in favor of non-monotonicity. There can be no question that the lag-CRP from, for instance the Howard et al. (2007) immediate free recall lag-CRP data, which consists of more than 7,000 trials, is more consistent with TCM when end-of-list context persists as part of the retrieval cue than when it does not (compare Figure 1a with Figure 3c). Farrell and Lewandowsky (in press) demonstrated that their two-parameter TCM_{evo} outperforms their two-parameter TCM_{pub} for all 14 of the data sets they considered, and often by a large margin, suggesting that the lag-CRP curves more closely resemble those predicted by their two-parameter TCM_{evo}. This test is not definitive either, though, because there may be other parameters that could be allowed to vary that would change the results. Perhaps a better approach would be to compare the model with ρ_{study} and ρ_{test} allowed to vary freely with the nested model in which they are constrained to have the same value.

It is also desirable, however, to be able to measure the properties of the lag-CRP in a model-independent way. It is actually quite difficult to measure persistent marginal tendency to recall items at various serial positions using the lag-CRP. Transitions to the first item, say, are isolated only at the most extreme negative lag. At other lags, they are obscured in the lag-CRP by transitions to other serial positions. Howard et al. (2008) attempted to control for persistent recency effects in their across-list CRP by reporting transitions relative to the distribution from a shuffled data set. This is also not an entirely satisfactory solution—if one were interested in persistent marginal serial position effects, it would be better to measure them directly.

One can imagine a two-dimensional CRP in which one keeps track of all transitions from each serial position to each other serial position, as in Figure 7 of Farrell and Lewandowsky (in press). The standard lag-CRP can be thought of as averaging along the diagonal of this transition matrix. If our interest is in persistent serial position effects we can calculate an end-justified CRP curve, as in Figure 2 of Farrell and Lewandowsky (in press). One could also invert the conditionality of the CRP and examine the distribution of lags to an item at a particular serial position, given that it was recalled. By comparing the distribution of lags to a particular serial position to the distribution of lags to its neighbors, one may be able to establish persistent marginal serial position effects above and beyond those generated by the contiguity effect without fitting a particular function. We should caution that the combinatorics of such analyses require a great deal of data. Most free recall data sets do not contain enough observations to enable a clean measurement of the lag-CRP; the number of observations required to accurately describe these statistical tools will require considerably more data in practice.

Appendix 2: Detailed modeling methods

In TCM, the current state of context $\mathbf{t}_i$ is generated from the previous state of context $\mathbf{t}_{i-1}$ and the current input $\mathbf{t}_i^{\text{IN}}$ according to

$$\mathbf{t}_i = \rho_i \mathbf{t}_{i-1} + \beta \mathbf{t}_i^{\text{IN}}$$

where β is a free parameter and ρ_i is chosen such that the length of $\mathbf{t}_i$ is unity. Note that this implies that the rate of contextual drift depends on the amount and nature of the input vector $\mathbf{t}_i^{\text{IN}}$. We treated the asymptotic rate of drift $\rho := \sqrt{1-\beta^2}$ as the parameter.

In the modeling reported here we associated item i to the state of context that preceded it, $\mathbf{t}_{i-1}$ rather than $\mathbf{t}_i$, consistent with recent treatments of TCM (e.g. Howard et al., 2006; Rao & Howard, 2008; Sederberg et al., in press; Howard et al., revised).

The distractors for the retention interval in the case of delayed free recall and continuous-distractor free recall were implemented by adding a vector orthogonal to all preceding states of context weighted by β and multiplying the preceding context vector by ρ_D.

The parameter γ was used to weight the degree of contextual retrieval upon repetition of an item. Farrell and Lewandowsky (in press) did not allow γ to vary, consistent with what was done in Howard and Kahana (2002), but inconsistent with more recent treatments (Howard et al., 2006, 2005; Sederberg et al., in press). The parameterization of the weighting here is somewhat different from previous work. We describe an orthogonal vector $\mathbf{c}$ for each item in the list. The $\mathbf{c}$ vector for each item remains fixed throughout. Each item is also associated with an $\mathbf{h}$ vector. These two components combine to generate the input vector $\mathbf{t}^{\text{IN}}$. If item A is presented at time step i, then the input pattern at time step i is given by

$$\mathbf{t}_i^{\text{IN}} \propto (1-\gamma)\mathbf{c}_A + \gamma \mathbf{h}_A,$$

where the proportionality symbol indicates that $\mathbf{t}_i^{\text{IN}}$ is normalized to be of unit length before entering into the evolution equation above. Each item's $\mathbf{h}$ vector is initialized to zero, and then updated according to

$$\Delta \mathbf{h}_A = \mathbf{t}_{i-1},$$

again assuming that item A was presented at time step i (see also Rao & Howard, 2008; Howard et al., revised).

The model we used here is closely related to the TCM_{evo} model used by Farrell and Lewandowsky (in press). It does, however, differ in several respects. First, we associated each item i to the preceding state of context $\mathbf{t}_{i-1}$ rather than $\mathbf{t}_i$. Second, we used a different value for the effective duration of the retention interval and γ, both of which were fixed parameters in the fits Farrell and Lewandowsky (in press) conducted (our definition of γ is also slightly different from some previous treatments). Third, they fit the model separately to each subject in each experiment and then reported an averaged result weighted by the number of observations made for each subject (see supplemental methods Farrell &

Lewandowsky, in press). This averaging across different numbers of observations at each lag accounts for the abrupt non-linearities in the fits of both TCM and the empirical models in the Farrell and Lewandowsky (in press) paper. Instead, we generated one analytic curve per condition and compare it informally to the data.