Memory Without Consolidation: Temporal Distinctiveness Explains Retroactive Interference

Ullrich K. H. Ecker

School of Psychology, University of Western Australia

Gordon D. A. Brown

Department of Psychology, University of Warwick

Stephan Lewandowsky

School of Experimental Psychology, University of Bristol, and

School of Psychology, University of Western Australia

Word count: 7,235 (excluding abstract, references, tables, and figure captions)

Author Note: Preparation of this paper was facilitated by a Discovery Grant and a Postdoctoral Fellowship from the Australian Research Council to UE, a Wolfson Research Merit Award from the Royal Society to SL, and UE's participation in the Swiss National Science Foundation's 2012 Summer School in Computational Modeling of Cognition. We thank Stewart Craig and Devon Spaapen for research assistance, and John Wixted for comments on an earlier draft. Address correspondence to the first author at the School of Psychology (M304), UWA, Crawley 6009, Australia. Telephone: +618 6488 3257; E-mail: ullrich.ecker@uwa.edu.au. The lab's web site is located at www.cogsciwa.com.

Abstract

Is consolidation needed to account for retroactive interference in free recall? Interfering mental activity during the retention interval of a memory task impairs performance, in particular if the interference occurs in temporal proximity to the encoding of the to-be-remembered (TBR) information. There are at least two rival theoretical accounts of this temporal gradient of retroactive interference. The cognitive neuroscience literature has suggested neural consolidation is a pivotal factor determining item recall. According to this account, interfering activity interrupts consolidation processes that would otherwise stabilize the memory representations of TBR items post-encoding. Temporal distinctiveness theory, by contrast, proposes that the retrievability of items depends on their isolation in psychological time. According to this theory, information processed after the encoding of TBR material will reduce the temporal distinctiveness of the TBR information. To test between these accounts, implementations of consolidation were added to the SIMPLE model of memory and learning. We report data from two experiments utilizing a two-list free recall paradigm. Modeling results imply that SIMPLE was able to model the data and did not benefit from the addition of consolidation. It is concluded that the temporal gradient of retroactive interference cannot be taken as evidence for memory consolidation.

Keywords: temporal distinctiveness; consolidation; SIMPLE; free recall; retroactive interference

Memory Without Consolidation: Temporal Distinctiveness Explains Retroactive Interference

Why do we remember the things we remember and forget the things we forget? Here we focus on—and provide evidence against—a central role for consolidation in memory for lists of unrelated material (free recall). One particularly prominent empirical finding is that memory for to-be-remembered (TBR) material (e.g., list L1 in Fig. 1) is impaired by interfering activity (e.g., study of a second list of items; L2 in Fig. 1) during the retention interval. Impairment is largest if interfering activity immediately follows learning (panel a in Fig. 1), and the bigger the temporal gap after learning, the smaller the impairment (panel b in Fig. 1). This phenomenon is known as the temporal gradient of retroactive interference (TGRI). It was first described by Müller and Pilzecker (1900) and was reviewed in Wixted (2004).

The cognitive neuroscience literature has proposed that neural consolidation is a pivotal factor determining item recall, and it has taken the existence of the TGRI as evidence for consolidation (Izquierdo, Schröder, Netto, & Medina, 1999; McGaugh, 2000; Wixted, 2004). Consolidation is a presumed neural process, acting over time—in particular during sleep or periods of low mental activity—that strengthens memories after they have been first formed, making them increasingly resistant to forgetting. Hence, consolidation is seen to counteract forgetting, and we forget if consolidation is interrupted.

According to this account, interfering activity will interrupt consolidation processes that stabilize memory representations of TBR items post-encoding. Consolidation theory thus assumes that the TGRI occurs because the larger the temporal gap between encoding of TBR material and the interrupting activity, the more consolidation can take place (Dewar, Cowan, & Della Sala, 2007; Wixted, 2004). This is because consolidation is assumed to have a fixed time-scale and hence the interrupting effect of additional study material should decrease with absolute time (i.e., the interval between L1 and L2). Performance may drop again when the interrupting material immediately precedes recall (panel c in Fig. 1), which sometimes leads to the assumption of an additional second process such as output interference.

However, current cognitive models of memory almost invariably accord no role to consolidation, assuming instead that forgetting results from cognitive interference, meaning that cognitive processing of additional information before or after learning can disrupt access to stored TBR information. Notwithstanding the lack of a consolidation mechanism, current cognitive memory models (e.g., Brown, Neath, & Chater, 2007; Farrell, 2012; Howard & Kahana, 2002; Lewandowsky & Farrell, 2008; Oberauer & Kliegl, 2006; Oberauer, 2009; Shiffrin & Steyvers, 1997) can accommodate a wide range of behavioral data. Here we examine whether the TGRI can also be accommodated without the assumption of consolidation.

One such cognitive model—the Scale-Invariant Memory, Perception, and LEarning model SIMPLE (Brown, Neath, & Chater, 2007)—implements a temporal distinctiveness theory. Temporal distinctiveness theory proposes that the discriminability and thus retrievability of items is a direct function of their isolation in time; information that is processed in temporal proximity to TBR information will reduce the temporal distinctiveness of the TBR information and hence impede its retrieval (Bjork & Whitten, 1974; Brown, Neath, & Chater, 2007; Brown, Chater, & Neath, 2008; Brown, Vousden, & McCormack, 2009; Brown & Lewandowsky, 2010; Glenberg & Swanson, 1986; Lewandowsky, Nimmo, & Brown, 2008; Nairne, Neath, Serra, & Byun, 1997; Neath & Crowder, 1990; Unsworth, Heitz, & Parks, 2008). According to this theory, information processed after the encoding of TBR material will compete with the TBR information at the time of retrieval and will thus potentially lead to interference-based forgetting of the TBR information. The principle of temporal distinctiveness predicts that relative time—rather than absolute time—is crucial: Specifically, in SIMPLE the discriminability of any two items (or sets of items, such as word lists) in memory will depend on the ratio of the temporal distances of the competing memory items at the time of retrieval. For example, an item A encoded 20 seconds ago will be easier to retrieve if an interfering item B was encoded 10 seconds ago (the ratio of temporal distances will be 2 : 1) compared to a situation where item B was encoded 15 seconds ago (the ratio of temporal distances will be 4 : 3). In other words, items will be better remembered if they are temporally isolated during presentation and hence are more discriminable. Thus the TGRI could emerge not because an increasing temporal gap allows for more consolidation, but because of the greater temporal distinctiveness conferred on TBR (i.e., L1) material (cf. Brown & Lewandowsky, 2010; Lewandowsky, Ecker, Farrell, & Brown, 2012).

Given these two different accounts of the TGRI, this study investigated whether the predictions of SIMPLE in a two-list free recall paradigm could be improved by adding a consolidation mechanism to the model, or whether temporal distinctiveness, by itself, is sufficient to account for the TGRI. In the following, we will first describe the basic assumptions and the standard computational implementation of SIMPLE, before we detail the two-list free recall task methodology employed in two experiments. When discussing the results of these experiments, we further specify the SIMPLE modeling applied to those data, and in particular the implementation of a consolidation mechanism into SIMPLE.

SIMPLE

SIMPLE assumes that each item encoded into memory is characterized by its position in a psychological memory space. In its most basic version—as outlined above—SIMPLE considers only one, temporal, dimension of this space, but in principle, the psychological space is considered multi-dimensional and might include other dimensions such as list context or an item's position on a list or in a spatial array. The basic temporal dimension is thought to reflect psychological, rather than absolute, time. This means that time is psychologically compressed as it passes, such that items appear closer together on the temporal dimension as time progresses, similar to equidistant telephone poles appearing closer together in perception as one moves away from them (Crowder, 1976). This means that the ratio of temporal distances discussed above will be expressed mathematically as a simple distance in log-space (e.g., a ratio of $2:1$ can be expressed as $log(2) - log(1)$). This notion of distance-ratios and log-distances illustrates one of the central aspects of the model, namely its scale-invariance. All other things being equal, presentation of items 20 and 10 seconds ago should lead to the same outcome as items presented 2 seconds and 1 second ago (because $20:10 = 2:1$ and $log(20) - log(10) = log(2) - log(1)$).

Formally, SIMPLE defines the similarity η of two items in memory as an exponential function of the psychological distance d between them,

$$\eta_{i,j} = e^{-cd_{i,j}}$$

where the parameter c defines how similarity reduces with distance and thus determines the 'sharpness' of memory representation. This is SIMPLE's main free parameter, and the larger c, the better memory performance.

The discriminability D of an item's memory trace is defined as a function of the summed similarity of an item to all other items,

$$D_i = \frac{1}{\sum\limits_{j=1}^{n} \eta_{i,j}}.$$

The recall probability P of an item is then a direct function of an item's discriminability,

$$P_i = \frac{1}{1 + e^{-s(D_i - t)}}$$

where t is a recall threshold (meaning that items below threshold-discriminability are omitted), and s is the slope of the transforming function determining the noisiness of the threshold.

A further computational step is required to accommodate data from a free recall task, where items can be recalled in any position, so as to integrate recall probabilities over all possible output positions. Following Lee and Pooley (2013), total recall probability is thus given by

$$\theta_i = 1 - \prod_{j=1}^{n} (1 - P_{i,j}).$$

Finally, in the present study we implemented an optional output interference mechanism to account for the well-established phenomenon that the active recall of each item during test will interfere with retrieval of all further items (Lewandowsky & Murdock, 1989; Tulving & Arbuckle, 1966). We thus added an output interference parameter o that reduced c for each output position,

$$c = c \times o^{i-1}.$$

This basic implementation of SIMPLE—with some versions augmented by a consolidation mechanism (see below for more details)—was used to model the data from two free recall experiments involving memorization of two consecutive word lists. In order to investigate the roles of absolute post-encoding time and relative time, we manipulated both the interval between the main TBR list (list 1) and the interfering list 2, as well as the interval between list 2 and list-1 recall. To reiterate, the first (L1-L2) interval should be most important from a consolidation point of view, while the ratio of the latter interval (L2-T) and the combined intervals (L1-T) should be the main determinant of recall from a temporal distinctiveness point of view.

Experiment 1

Experiment 1 involved the study of two unrelated word lists, with a simple and easy tone-detection distractor task of varying duration given between the two study lists and between study of the second list and the final free recall test. A tone task was used—rather than complete rest—to have some control over participants' behavior and to prevent rehearsal; we assumed the tone task to be sufficiently easy that consolidation would not be prevented. The free recall test always targeted recall of the first list, although subsequent recall of the second list was required on a random 50% of trials, in order to maintain the study intention for both lists.

More specifically, the experiment used a within-subjects design. There were four experimental conditions, created by fully crossing the duration of the two distractor task intervals in each trial. As illustrated in Fig. 2, distractor-task duration was either short (60 s) or long (240 s), resulting in the conditions short-short (SS), short-long (SL), long-short (LS), and long-long (LL). Each participant completed two experimental sessions on separate days, with each condition being presented twice per session (i.e., 16 trials in total). Condition order was randomly determined within each session for each participant.

We derived some predictions about mean recall accuracy from temporal distinctiveness theory. With the simplifying assumption that all list items were presented at the same point in time—that is, considering only the intervals between study/test events—temporal distinctiveness theory predicted that recall should be best in the LS condition and worst in the SL condition, because at the time of recall list 1 had the clearest temporal isolation in condition LS and the lowest temporal isolation in condition SL. We further hypothesized that conditions SS and LL should produce intermediate and similar levels of recall, based on the time-scale invariance assumption of temporal distinctiveness theory (i.e., SL $<$ SS $=$ LL $<$ LS). In contrast, a strict version of consolidation theory would predict that the main factor for recall performance should be the amount of uninterrupted time following list-1 encoding, and would hence predict that the LS and LL conditions should be similar to each other and superior to SS and SL conditions (i.e., SL $=$ SS $<$ LL $=$ LS).

Given the minimum retention interval of 130 seconds, we further predicted that serial position curves would show a strong primacy effect but only a small or absent recency effect (cf. Bjork & Whitten, 1974). We predicted that SIMPLE would be able to capture both the overall shape of the serial-position curve and the expected differentiation between conditions.

Methods

Participants. The participants were 23 members of the University of Western Australia campus community (15 females, 8 males; mean age 19.3 years, age range 17-26 years). Participants received either course credit or AU$20 remuneration for their participation in the two one-hour sessions.

Apparatus. The experiment was undertaken on a Windows based computer using a MATLAB program designed with the aid of the Psychophysics toolbox (Brainard, 1997).

Stimuli. Words were selected from the MRC Psycholinguistic database (available online at `http://www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm`). We selected 320 words from the database using the following constraints: words were (a) one syllable nouns, (b) between 3 and 6 letters, (c) had a Kucera-Francis frequency of greater than 28, (d) had familiarity and concreteness ratings of at least 400, and (e) were not confusable with other words on the list (e.g., 'male' was removed because of its similarity to 'mail'). A total of 32 10-word lists were generated for each participant. Each list was created by taking a random word from the first tenth of the full, alphabetically ordered, list, then a word from the second tenth of the list, and so on. All lists were generated in the same manner, without replacement of words. Additionally, list generation was constrained such that adjacent words in the full list would not appear in either of the two word lists within a single trial. These constraints were placed on list generation to reduce the chance that phonetically similar words would appear in the same trial. Word lists for practice trials were also taken from the MRC Psycholinguistic database, using the same constraints given above, but with a Kucera-Francis frequency of 25 to 28.

Procedure. In each experimental trial, participants were required to memorize two word lists with 10 words each. Additionally, they completed two blocks of the tone-detection distractor task. Participants were instructed not to rehearse the word lists during the distractor task, and to aim for 100% accuracy in the tone-detection task, which is described in detail below. Each experimental trial involved five stages: (a) the presentation of to-be-remembered word list 1, (b) the first block of the tone task, (c) the presentation of to-be-remembered word list 2, (d) the second block of the tone task, and (e) word list recall, targeting word list 1 on all trials, and additionally targeting word list 2 on half the trials.

At the start of each trial, the text "Prepare for word list" was presented centrally for 1 s, followed by the presentation of a fixation cross for 500 ms. The first 10-word list was then presented. Words were centrally presented one at a time, for 500 ms each, with an inter-stimulus interval (ISI) of 400 ms. The words were presented in all capital letters in black on a white background.

After the presentation of the word list, the first tone task commenced. The text "Tone task (Press 'h' for high tones)" appeared on the screen for 750 ms, presented in white on a black background. A central fixation cross was then displayed and a sequence of tones was presented via headphones. Each tone was played for 150 ms, with a 600 ms delay between tones. Two pitches of tones were played: approximately 80% of tones were low tones (440 Hz; i.e., note A4) and approximately 20% of tones were high tones (523 Hz; i.e., note C5). The order of tones was randomly determined, with the constraint that the first tone was always a standard low tone. Participants were instructed to press the 'h' key as quickly as possible upon hearing a high tone. If, on the presentation of a high tone, a response was not recorded within 600 ms from tone onset, the word "MISS" was presented in red for 150 ms. If the 'h' key was pressed on the presentation of a low tone, the words "FALSE ALARM" were immediately presented in red text, until the presentation of the next tone. A tone trial was recorded as correct if the participant responded to a high tone, or provided no response to a low tone. The duration of the tone task was either short (60 s, corresponding to 79 tones) or long (240 s, corresponding to 319 tones), depending on the condition. After the first tone-task block, a second word list was presented in the same manner as the first, immediately followed by a second tone task, which again was either short or long, depending on condition.

Following the second tone task, participants were prompted to recall the first word list, with the text "Please recall the FIRST word list—Start speaking after the tone" presented in black on a white screen. After 2 s, a 450 ms tone (392 Hz; i.e., G4) was played to signal that recording had started. Participants provided verbal recall, which was recorded via a microphone built into the headphones. Maximum time for recall was 30 s, but participants could end the recall test at any time after 15 s, when the text "Press 's' if you cannot recall any more words" was presented on screen. On half the trials, participants were then prompted to recall the second word list, in the same manner. The trials on which second-list recall was required were randomly determined, with the constraint that there would be no more than two trials in a row without recall of the second list.

Participants completed one practice trial at the start of each of the two sessions. The practice trials were identical to experimental trials, except that the distractor task intervals were only 10 s. Participants were always prompted to recall both word lists on the practice trials.

Results

Mean recall performance in the four experimental conditions is shown in Fig. 3. A 2 $\times$ 2 repeated measures ANOVA yielded significant main effects of both the L1-L2 interval, $F(1, 22) = 8.74, MSE = 0.006, p < .01, \eta_p^2 = .28$, and the L2-T interval, $F(1, 22) = 5.97, MSE = 0.007, p = .02, \eta_p^2 = .21$, but no interaction, $F < 1$. In line with the predictions of temporal distinctiveness theory, it was thus not only the L1-L2 interval that affected performance, but also the L2-T interval. Moreover, the performance order of conditions—SL $<$ SS $=$ LL $<$ LS—was exactly as predicted by temporal distinctiveness theory (the SL vs. SS and LL vs. LS effects were both marginally significant in direct contrasts, $F(1, 22) = 3.09, p = .09$, and $F(1, 22) = 3.76, p = .06$).

This analysis at the level of means suggests that temporal distinctiveness theory can handle the data of Experiment 1 well. However, it is still possible that consolidation makes a significant contribution to word-list recall. To test whether a consolidation mechanism is necessary to (more fully) account for the data of Experiment 1, we used a computational modeling approach. The modeling was performed on the serial position curve data, which are shown in Fig. 4. We applied a large number (30+) of model versions. First, we implemented two basic, consolidation-free versions of SIMPLE. One version was the basic version described above, with only one, temporal dimension. The second version added a second dimension of psychological space, namely list context. This allowed the model to increase the psychological distance between list-1 and list-2 item representations on a dimension other than time. The dimensions were either equally weighted, or a free weighting parameter w was additionally introduced.

We then added various implementations of a consolidation mechanism to these basic models. It was required to consider a large number of consolidation implementations because consolidation as a computational process is underspecified in the literature (cf. Ecker & Lewandowsky, 2012). All implementations shared one characteristic: They increased the c parameter of SIMPLE over absolute time post-encoding of L1, to increase the distinctiveness of items in memory. However, we varied (1) the exact starting point—we let consolidation begin immediately upon encoding of each item, or delayed consolidation onset until the end of list presentation, (2) the exact end-point—we let consolidation run only until the onset of the disruptive processing of L2, or let it re-start afterwards, and (3) the exact shape—we let c increase linearly or nonlinearly.

More specifically, a linear increase was modeled either by letting c grow by a fixed amount α for each unit of absolute time,

$$c_{consol} = c + \alpha \times time$$

or by letting c grow by a fixed proportion of its initial value for each unit of time,

$$c_{consol} = c + \alpha c \times time.$$

A nonlinear increase was modeled by adding an additional exponent parameter b,

$$c_{consol} = c + \alpha c \times time^b.$$

This power function allowed for either accelerating growth of c over time (in case of $b > 1$)—that is, an implementation whereby the impact of consolidation grows over time—or (in case of $0 < b < 1$) for decelerating growth—that is, an implementation whereby the impact of consolidation is strong initially but reduces over time. Based on previous research and logical constraints, we limited the range of parameters as summarized in Table 1.

We fit the models to the serial position curve data for each participant separately using maximum likelihood estimation, utilizing MATLAB's $fminsearch$ algorithm. We used deviance and both BIC and $AICc$ to gauge model fit and select the best-fitting models after taking account of the number of free model parameters. As far as the consolidation models are concerned, models implementing consolidation with a starting point at the end of list 1, and an end-point at the onset of list 2, outperformed the other consolidation models. As would be expected, the most complex consolidation model—the two-dimensional model with nonlinear (mostly exponentially growing) consolidation—had the lowest overall deviance.

We then selected the best 5-, 6-, and 7-parameter consolidation models, and used BIC and $AICc$ weights to adjudicate between the consolidation-free SIMPLE models and the best consolidation models. Results are summarized in Table 2, and show that the improvement in model fit (i.e., the deviance reduction) associated with adding consolidation was too marginal to justify the increase in complexity from adding additional parameters. In other words, the BIC and $AICc$ weights clearly favored the 4-parameter, consolidation-free SIMPLE models.

More specifically, one of the two 4-parameter consolidation-free models achieved a better fit than any of the consolidation models—as indicated by an $AICc$ difference of 3 or more (cf. Raftery, 1996)—in 10 out of the 23 participants. Using the BIC difference as an indicator (again, accepting a difference of 3 or greater as evidence for the model with the lower criterion value), the consolidation-free models achieved a better fit in 17 out of the 23 participants. In no case did one of the consolidation models achieve a better fit than the 4-parameter models with no consolidation.[1]

The considerable flexibility of SIMPLE allowed us to then compare the models discussed so far with versions of SIMPLE that featured a consolidation mechanism but no temporal dimension. Following precedent (Lewandowsky, Duncan, & Brown, 2004), we implemented item list-position as an additional dimension into SIMPLE to replace the temporal dimension. We then ran another model comparison, pitting the best consolidation-free model against the best model combining temporal distinctiveness and consolidation, and the best time-free consolidation model. Results of the model comparison, as well as the mean parameter estimates for each of the three models, are summarized in Table 3. The model comparison established the consolidation-free SIMPLE model as the best-fitting model when number of parameters is taken into account.

More specifically, the consolidation-free SIMPLE model was the only model to achieve a superior fit to the data than *both* the other two models for a subset of participants. Utilizing the $AICc$ difference as an indicator (again, accepting a difference of 3 or greater as evidence for the model with the lower criterion value), the consolidation-free model achieved a better fit in 3 participants; using the BIC difference as an indicator, the consolidation-free model achieved a superior fit in 8 participants. In no case did one of the consolidation models achieve a better fit than the consolidation-free model.

Discussion

Consolidation theory predicts that in a two-list free recall paradigm, recall of the first list (L1) will be a direct function of the absolute amount of uninterrupted time available post L1 encoding. Temporal distinctiveness theory, by contrast, predicts that recall of the

[1]The deviances, $AICc$ and BIC values, and parameter estimates of all participants and all models featured in Tables 2, 3, 4, and 5 are available from the first author on request.

first list will be a direct function of the ratio of the retention intervals of both lists L1 and L2 (i.e., the ratio of L1-T and L2-T intervals): Specifically, L1 recall will be good if the ratio is large, and bad if the ratio is small. To test this prediction, Experiment 1 manipulated the temporal distance between two TBR lists of words (the L1-L2 interval), and the temporal distance between the interfering second list and test (the L2-T interval).

The pattern of means confirmed the predictions of temporal distinctiveness theory. Recall was best when the L1-L2 interval was long and the L2-T interval short—implying that the ratio of the L1-T and the L2-T intervals was large—and worst when the opposite was true, that is, when the L1-L2 interval was short and the L2-T interval long—implying that the ratio of the L1-T and the L2-T intervals was small. In other words, the more the first list was isolated in time, the better its retrieval. Moreover, the fact that the two conditions with an equal ratio of L1-T and L2-T intervals (i.e., the short-short and the long-long conditions) produced the same performance level despite large differences in absolute retention interval supports the scale-invariance notion of temporal distinctiveness theory in general and SIMPLE in particular.

Modeling the serial position curves of Experiment 1 produced further evidence that temporal distinctiveness theory is well-equipped to explain the present set of data. We implemented both one-dimensional (temporal distinctiveness only) and two-dimensional (temporal distinctiveness plus list context) versions of SIMPLE, and then added a large variety of potential consolidation mechanisms. Additionally, we implemented a two-dimensional SIMPLE version free of temporal distinctiveness, with list-context and item-position dimensions only, augmented by consolidation. Model comparisons showed that a two-dimensional consolidation-free version of SIMPLE was the preferred model. The strength of evidence for this model was not statistically overwhelming—it was preferred over the other models in the final 3-model comparison by a factor of about 3 to 4—but in terms of our main question, namely whether or not the SIMPLE predictions could be improved by the addition of a consolidation mechanism, the model comparison certainly did not suggest the need for inclusion of a consolidation mechanism to improve SIMPLE's fit.

Experiment 2

Experiment 2 was similar to Experiment 1 and was conducted to address the possible concern that the tone-detection task in Experiment 1 might have allowed sub-vocal rehearsal. It is not entirely clear what the impact of rehearsal on the outcome of Experiment 1 and the modeling would have been, but it is a possibility that intentional rehearsal might have masked the effects of consolidation in Experiment 1, for example by occurring selectively for some (early or late) items but not others. In Experiment 2, the tone-detection distractor task was thus replaced with a number task, which required participants to verbalize a series of two-digit numbers presented on screen. The number task was designed to prevent word-list rehearsal via articulatory suppression, while still being easy enough so as not to interrupt consolidation. Numbers were used rather than words or text to minimize specific interference with the words on the list.

Method

Participants. The participants were 25 members of the University of Western Australia campus community (19 females, 6 males; mean age 19.8 years; age range 18-31 years). Participants received either course credit or AU$20 remuneration for their participation.

Apparatus, Design, Stimuli, and Procedure. The apparatus, design, stimuli, and procedure were identical to Experiment 1, with the following exceptions. The distractor task was a number-reading task, which required participants to read out loud a series of two-digit numbers presented individually on the screen. Each number in the sequence was randomly selected from a uniform distribution ranging from 10 to 99.

The task began with the text "Number task (Say each number that appears on the screen)" presented centrally on the screen for 750 ms. A sequence of two digit numbers was then presented on the screen for 1 s each, with a 250 ms ISI. All text was presented in dark gray on a light grey background. Participants were instructed to read aloud each number that appeared on the screen. Utterances were recorded and compliance checked via inspection of the sound files. A total of 47 or 191 numbers were presented in the short and long intervals, respectively.

Results

Mean recall performance in the four experimental conditions of Experiment 2 is shown in Fig. 5. A 2 × 2 repeated measures ANOVA yielded only a marginally significant main effect of the L2-T interval, $F(1, 24) = 3.04, MSE = 0.004, p = .09, \eta_p^2 = .11$ (other effects, $F < 1.02$). Compared to Experiment 1, we replicated the SS vs. SL difference (again, this was marginally significant, $F(1, 24) = 3.80, p = .06$) but the long L1-L2 interval data differed from Experiment 1.

Despite the relative absence of mean condition differences, modeling these data could be informative, for example regarding the question of whether SIMPLE would be able to handle the superiority of the SS condition, in particular in the primacy region of the serial position curve. Serial position curve data from Experiment 2 are shown in Fig. 6.

Again, the most complex consolidation model—the two-dimensional model with nonlinear consolidation—had the lowest overall deviance. In contrast to Experiment 1, however, consolidation was mainly estimated to lead to decelerated, logarithmic-like growth in memory strength, rather than exponential growth.

As in the Experiment 1 analyses, we selected the best 5-, 6-, and 7-parameter consolidation models, and used BIC and $AICc$ weights to adjudicate between the consolidation-free SIMPLE models and the best consolidation models. Results are summarized in Table 4, and show that the improvement in model fit (i.e., the deviance reduction) associated with adding consolidation was too marginal to justify the increase in complexity from adding additional parameters. In other words, the BIC and $AICc$ weights clearly favored the 4-parameter, consolidation-free SIMPLE models.

More specifically, one of the two 4-parameter consolidation-free models achieved a better fit to the data than any of the consolidation models—as indicated by an $AICc$ difference of 3 or more—in 4 out of the 25 participants. Using a BIC difference greater than 3 as an indicator, the consolidation-free models achieved a better fit in 21 out of the 25 participants. In no case did one of the consolidation models achieve a better fit than the 4-parameter models with no consolidation.

In an additional model comparison, we then compared the best consolidation-free model with the best model combining temporal distinctiveness and consolidation, and the best time-free model. Results as well as mean parameter estimates are summarized in Table 5. As in Experiment 1, the model comparison established the consolidation-free SIMPLE model to be the best-fitting model.

More specifically, the consolidation-free SIMPLE model achieved a superior fit than *both* the other two models for a subset of participants. Utilizing the $AICc$ difference greater than 3 as an indicator, the consolidation-free model achieved a better fit in 3 participants; using the BIC difference as an indicator, the consolidation-free model achieved a superior fit in 15 participants. The time-free consolidation model achieved a better fit than the consolidation-free model in one case.

Discussion

In contrast to Experiment 1, Experiment 2 did not produce much variation between conditions. In particular, the long L1-L2 interval data differed from Experiment 1: accuracy was lower, and there was little difference between the two long L1-L2 interval conditions. The reason for this is likely the more demanding distractor task, which over the course of four minutes might have produced significant interference or disruption of consolidation. Nonetheless, the effect of the L2-T interval—albeit weak—is in line with temporal distinctiveness theory, and the complete absence of an L1-L2 interval effect is difficult to reconcile with the predictions of consolidation theory.

Modeling the serial position curves of Experiment 2 led to the same conclusion as the modeling of Experiment 1. The comparison of the best consolidation-free model with the best temporal-distinctiveness-plus-consolidation model and the best temporal-distinctiveness-free consolidation model revealed that a two-dimensional consolidation-free version of SIMPLE was the preferred model.

Again, the strength of evidence for this model was not overwhelming—it was preferred over the other models in the comparison by a factor of about 3 to 5—but in terms of our question whether or not the SIMPLE predictions could be improved by the addition of a consolidation mechanism, the model comparison again did not suggest the inclusion of a consolidation mechanism to improve SIMPLE's fit.

General Discussion

We presented two experiments demonstrating that temporal distinctiveness theory (Bjork & Whitten, 1974; Brown, Neath, & Chater, 2007; Glenberg & Swanson, 1986) can account parsimoniously for delayed free-recall data from a two-list paradigm without the assumption of consolidation. We modeled serial position curves using a computational instantiation of temporal distinctiveness theory, SIMPLE. Our efforts to improve SIMPLE's approximation of the data by implementing a consolidation mechanism proved unsuccessful, and hence the present study does not support the inclusion of a consolidation mechanism into SIMPLE. In the General Discussion we first consider possible alternative interpretations of the data, then discuss the relationship between high-level temporal distinctiveness models such as SIMPLE and their possible neurocomputational underpinnings.

Interpretations of Results

Our design and analysis have tacitly assumed that what would interfere with the consolidation of memory for one list of items is primarily the presentation of a second list of to-be-remembered items. However, it has also been suggested that the process of forming *any* new memory representation can disrupt consolidation (e.g., Wixted, 2004; Mednick, Cai, Shuman, Anagnostaras, & Wixted, 2011). It is therefore theoretically possible that our rehearsal-prevention task (tone detection during the gaps between lists) itself led to the formation of new memories, thus preventing or reducing consolidation during the temporal gaps.

Although it may be difficult to exclude such a possibility entirely, we concluded that a temporal distinctiveness account of the present data is preferable for at least three reasons. First, both informal observation from our own laboratories and published work suggests that post-list memory for distractors is either virtually absent or at least small compared with memory for target items (e.g. McFarlane & Humphreys, 2012). While participants undoubtedly retain a memory of having performed the tone task, just as they retain a memory of having viewed some word lists, such memories seem very different from memory for specific items and their formation seem less likely to interfere with any plausible process-level consolidation.

Second, we note that a basic memory of having performed the tone task (and of the tones themselves) would have been formed in all conditions, in that the long intervals of tone detection merely appended additional time on the same activity. It is known from previous work that extending a distractor task of uniform stimuli does nothing to cause further forgetting (Lewandowsky, Geiger, & Oberauer, 2008).

Third, there was a clear overall effect of the L1-L2 gap on memory for List 1, and a consolidation account could not explain this effect if the tone detection task led to a significant amount of consolidation-preventing memory formation. Thus—at least in the present paradigm—if the tone detection task prevented consolidation, then positive effects on memory of post-list temporal gaps could not be interpreted as evidence for consolidation, and consolidation models could not be applied to experiments examining memory for lists when rehearsal is prevented. This would at least severely limit the applicability of consolidation models. In other words, while we cannot exclude the possibility that formation of implicit memories (e.g., of the tones presented in the rehearsal-preventing task) will impede consolidation, if consolidation were so sensitive to disruption then other types of evidence cited in support of consolidation would need alternative interpretation, as very little information would ever be consolidated unless a person sleeps. Other consolidation theorists have, however, argued that periods of wakeful rest following encoding benefit memory due to uninterrupted consolidation (e.g., Dewar, Alber, Butler, Cowan, & Della Sala, 2012), in a scenario where participants would have likewise formed some basic memory for the rest-period episode. Any argument that consolidation can survive some types of ongoing activity but not others risks unfalsifiability in the absence of a sufficiently detailed specification of the conditions under which consolidation is possible.

Recent evidence suggests that reactivation of a memory trace—whether externally cued, voluntary, or spontaneous—might be consolidation's crucial "method of action" (e.g., Oudiette & Paller, 2013; Staresina, Alink, Kriegeskorte, & Henson, 2013). If reactivation is the mechanism by which consolidation is implemented, then it could be argued that the reason a delay between lists affects recall of the first list is that, after presentation of both lists, two lists compete for reactivation. However, if consolidation (interpreted as reactivation) strengthens memory traces only or mainly in the interval between the two lists (i.e., before L2 competes for reactivation resources), then one of our implementations of consolidation (e.g., where strengthening occurs only in the L1-L2 interval, or where strengthening continues until test but its effect reduces over time viz. a nonlinear model with $b < 1$) should have captured this.

We therefore conclude that our evidence questions one key source of behavioral evidence for consolidation—the TGRI. A similar approach may explain some—but not all—other evidence for consolidation in memory. For example, the beneficial effects of post-encoding sleep on recall can easily be explained within a temporal distinctiveness framework—periods of sleep will render information acquired before sleep onset more isolated in time and will thus reduce retroactive interference. However, we do not claim to address all evidence for the existence of consolidation. For example, it is unclear how differential effects of specific sleep *phases* could be explained without recourse to a specific trace-strengthening process such as consolidation. In particular, it seems that declarative memory benefits specifically from periods of slow-wave sleep but not other forms of sleep (Born, Rasch, & Gais, 2006; Born, 2010). Support for this notion comes from studies demonstrating memory-boosting effects of fronto-cortical stimulation (e.g., transcranial direct current stimulation) during slow-wave sleep (Marshall, Mölle, Hallschmid, & Born, 2004; Marshall, Helgadottir, Mölle, & Born, 2006). Other studies have demonstrated that cuing of recently encoded information during slow-wave sleep (but not other sleep phases) can improve later retrieval of that information. For example, Rasch, Büchel, Gais, and Born (2007) presented a scent during encoding of object-location associations, and also during a phase of post-encoding sleep. They found that the scent-cue boosted subsequent memory performance, but only when it was presented during slow-wave sleep.

Relation Between SIMPLE and Neuroscience

SIMPLE is a higher-level model that deliberately abstracts away from specific brain states and does not specify the neural processes underlying encoding and retrieval. It seems clear that both high-level models (to capture regularities in the behavioral data) and lower-level models (to capture the relevant neurocomputational processes) will both be necessary for a complete account; it is also important that accounts at different levels eventually be integrated as far as possible. Fortunately, there has recently been considerable progress towards understanding how neural hardware might underpin the types of scale-invariant temporal distinctiveness principles embodied in SIMPLE.

First, there is mounting evidence that the medial temporal lobe (MTL) involves representations that code for the time at which events occur (for a comprehensive recent review, see Howard & Eichenbaum, 2013). One suggestion is that the hippocampus may play a role very similar to that assumed by a number of recent models of human memory, and that temporal distinctiveness effects arise due to temporal/contextual representations associated with memories in MTL.

Specifically, several models assume that a gradually changing temporal context signal is a central component of a memory system (e.g., Brown, Preece, & Hulme, 2000; Howard & Kahana, 2002; Polyn, Norman, & Kahana, 2009). For example, the OSCAR model of Brown et al. (2000) assumes that item representations are associated to a gradually time-varying signal constructed from oscillators of different frequencies, that serially ordered recall occurs when the temporal context signal can be "replayed", and that the recency of a memory is represented by the similarity between the current state of the temporal context and the state of the temporal context associated with the relevant item.

A typical property of gradually changing temporal contexts like the one OSCAR uses is that states of the context are more similar when they represent nearby locations in time: the further apart the times, the more different are the states of the temporal context. Thus, if temporal context signals are used as retrieval cues, as in the OSCAR model, retrieval cues and hence the items they are associated with will be less confusable when the items, and the temporal context they were associated with, occurred far apart in time. This gives rise to temporal distinctiveness effects. Recency effects can also emerge when temporal contexts of learning cannot be completely reinstated at retrieval, due to contextual overlap between learning context and testing context. There is now considerable evidence, from both the human and animal literature, that a temporal context signal may be represented in MTL. For example, MacDonald, Lepage, Eden, and Eichenbaum (2011) investigated sequence learning in rats and reported evidence for an ensemble of cells that become sequentially activated post stimulus presentation, such that the ensemble in its entirety can code for the passage of time (e.g., the duration of a retention interval); the cells in these ensembles have thus been labeled "time cells" (also see Kraus, Robinson, White, Eichenbaum, & Hasselmo, 2013).

In humans, Howard, Viskontas, Shankar, and Fried (2012) recorded the activity of human MTL neurons and found evidence of a gradually changing vector of neural activity during a continuous recognition task. Moreover, the temporal signal appeared partially to return to a previous state (the state associated with the first presentation of an item) when a repeated item was presented, consistent with the idea that presentation of an item activates temporal contexts previously associated with that item. Using recordings from neuropsychological patients, Manning, Polyn, Baltuch, Litt, and Kahana (2011) found a neural signature of context reinstatement at retrieval. There is therefore considerable emerging neural evidence for a temporal-contextual signal of the type that is (a) an important component of many memory models and (b) possibly responsible for temporal distinctiveness and recency effects of the type captured at a higher level by SIMPLE. Although the various temporal context models differ in a number of important ways, it is encouraging that an important common feature (the temporal context signal itself) appears to reflect neural activity in the MTL. Moreover, it seems that a gradually time varying signal in the MTL will give rise to temporal distinctiveness effects to the extent that states of the temporal context signal are used as retrieval cues to distinguish items.

More recent models have begun to integrate the levels. For example, computational neuroscience work by Shankar and Howard (2012) and Howard and Eichenbaum (2013) has demonstrated that a network model consisting of three cell arrays—one representing incoming stimuli, one comprising computational "time cells", and an intermediate layer required for the computation of temporal history from the input—can code a *scale-invariant* representation of temporal item history. In this model, each incoming stimulus is associated with the recent temporal history of encountered stimuli, and the temporal representations of stimuli become less accurate as time elapses. This maps exactly onto SIMPLE's assumption that items in memory become less discriminable as time passes.

We also note that SIMPLE and the temporal distinctiveness notion can be meaningfully applied to data from neuropsychological patients with hippocampal damage.

For example, while the temporally graded memory loss found in retrograde amnesia has frequently been used as evidence for the importance of long-term consolidation, there exists an alternative conceptualization. SIMPLE (and also the Howard and Eichenbaum (2013) neurocomputational model) suggests that the temporal dimension becomes less useful and hence less important for the retrieval of memories the older these memories become—because temporal/contextual representations become ever less discriminable as they recede into the past. The discovery of hippocampal "time cell" arrays makes it plausible that damage to the hippocampus might lead to a relatively selective loss of access to the temporal dimension of mnemonic representations. Losing access to this temporal dimension would lead to an impairment in the retrieval of recent but not temporally distant memories, or in other words, to a temporally graded retrieval impairment for recent memories (for a simulation that demonstrates this point, see Brown & Lewandowsky, 2010). Such a result fits well with data and models suggesting that hippocampal amnesia can be understood in temporal distinctiveness terms, rather than as reflecting disruption of a specific memory system, when the abnormal pattern of rehearsal exhibited by amnesic patients is taken into account (Brown, Della Sala, Foster, & Vousden, 2007; Dewar, Brown, & Della Sala, 2011).

Concluding Remarks

We argue that for consolidation theory to have any serious impact on contemporary cognitive models of memory, it requires further specification (cf. Ecker & Lewandowsky, 2012; Lewandowsky et al., 2012). The reverse is also the case: If cognitive models of memory are to be taken seriously by neuroscientists, such models must explain effects that have been attributed to consolidation either by incorporating well-specified consolidation mechanisms or, as here, by arguing that specific patterns of data can be explained without it (similarly, for a temporal-context account of data taken as evidence for *re*consolidation, see Sederberg, Gershman, Polyn, & Norman, 2011).

The under-specification of consolidation presented a major challenge for the current modeling, making it necessary to implement several dozen variations of a consolidation mechanism in a tedious trial-and-error approach. The above-mentioned sleep research has taken important steps towards specification, but many questions have not yet received satisfactory answers: When exactly does the consolidation of an item begin? What determines the disruption of consolidation? What is the shape of its "strengthening curve"? What is its time course? As discussed by Ecker and Lewandowsky (2012), the under-specification of consolidation's time-scale(s) in particular is a serious issue because various types of consolidation have been proposed to account for data across various time-scales (cf. Dudai, 2004), such that any empirical result can be explained by the "right mixture" of forgetting and various kinds of consolidation. Despite our unsuccessful attempts to accord a role to consolidation in the present modeling, specification of these parameters will constrain future modeling work, such that integration of a consolidation mechanism might ultimately strengthen cognitive memory models. Until such specification, however, consolidation theory runs the risk of being diluted and "über-applied" to empirical evidence for which current cognitive models, such as SIMPLE, provide both quantifiable predictions and parsimonious explanations.

We conclude that there should be more emphasis on the integration of computational modeling and neuroscientific concepts. Behavioral evidence for time-scale invariance has presented a challenge to neuroscientific accounts of episodic memory, which is only beginning to be addressed. We argue that the evidence for a role of temporal distinctiveness suggests that neuroscientists should more seriously consider it as an additional relevant mechanism in their work on episodic memory.

References

Bjork, R., & Whitten, W. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, *6*, 173-189.

Born, J. (2010). Slow-wave sleep and the consolidation of long-term memory. *World Journal of Biological Psychiatry*, *11*, 16–21. doi: 10.3109/15622971003637637

Born, J., Rasch, B., & Gais, S. (2006). Sleep to remember. *The Neuroscientist*, *12*, 410–424.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433-436.

Brown, G. D. A., Chater, N., & Neath, I. (2008). Serial and free recall: Common effects and common mechanisms? A reply to Murdock (2008). *Psychological Review*, *115*, 781–785.

Brown, G. D. A., Della Sala, S., Foster, J. K., & Vousden, J. I. (2007). Amnesia, rehearsal, and temporal distinctiveness models of recall. *Psychonomic Bulletin & Review*, *14*, 256–260.

Brown, G. D. A., & Lewandowsky, S. (2010). Forgetting in memory models: Arguments against trace decay and consolidation failure. In S. Della Sala (Ed.), *Forgetting* (pp. 49–75). Hove, UK: Psychology Press.

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

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

Brown, G. D. A., Vousden, J. I., & McCormack, T. (2009). Memory retrieval as temporal discrimination. *Journal of Memory & Language*, *60*, 194–208.

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

Dewar, M. T., Alber, J., Butler, C., Cowan, N., & Della Sala, S. (2012). Brief wakeful resting boosts new memories over the long term. *Psychological Science*, *23*, 955-960.

Dewar, M. T., Brown, G. D. A., & Della Sala, S. (2011). Restoring primacy in amnesic free recall: Evidence for the recency theory of primacy. *Cognitive Neuropsychology*, *28*, 386-396.

Dewar, M. T., Cowan, N., & Della Sala, S. (2007). Forgetting due to retroactive interference: A fusion of Müller and Pilzecker's (1900) early insights into everyday forgetting and recent research on anterograde amnesia. *Cortex*, *43*, 616–634.

Dudai, Y. (2004). The neurobiology of consolidation, or, how stable is the engram? *Annual Review of Psychology*, *55*, 51–86.

Ecker, U. K. H., & Lewandowsky, S. (2012). Computational constraints in cognitive theories of forgetting. *Frontiers in Psychology*, *3*, 400.

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

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

Howard, M. W., & Eichenbaum, H. (2013). The hippocampus, time, and memory across scales. *Journal of Experimental Psychology: General*, *142*(4), 1211-1230.

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

Howard, M. W., Viskontas, I. V., Shankar, K. H., & Fried, I. (2012). Ensembles of human mtl neurons "jump back in time" in response to a repeated stimulus. *Hippocampus*, *22*, 1833-1847.

Izquierdo, I., Schröder, N., Netto, C. A., & Medina, J. H. (1999). Novelty causes time-dependent retrograde amnesia for one-trial avoidance in rats through nmda receptor- and camkii-dependent mechanisms in the hippocampus. *European Journal of Neuroscience*, *11*, 3323-3328.

Kraus, B., Robinson, R., White, J., Eichenbaum, H., & Hasselmo, M. (2013). Hippocampal "time cells": Time versus path integration. *Neuron*, *78*(6), 1090-1101.

Lee, M., & Pooley, J. (2013). Correcting the SIMPLE model of free recall. *Psychological Review*, *120*, 293-296.

Lewandowsky, S., Duncan, M., & Brown, G. D. A. (2004). Time does not cause forgetting in short-term serial recall. *Psychonomic Bulletin & Review*, *11*, 771-790.

Lewandowsky, S., Ecker, U. K. H., Farrell, S., & Brown, G. D. A. (2012). Models of cognition and constraints from neuroscience: A case study involving consolidation. *Australian Journal of Psychology*, *64*, 37-45.

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. In B. H. Ross (Ed.), *The psychology of learning and motivation, 49* (pp. 1–48). London, UK: Elsevier.

Lewandowsky, S., Geiger, S. M., & Oberauer, K. (2008). Interference-based forgetting in verbal short-term memory. *Journal of Memory and Language*, *59*, 200-222.

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

Lewandowsky, S., Nimmo, L. M., & Brown, G. D. A. (2008). When temporal isolation benefits memory for serial order. *Journal of Memory and Language*, *58*, 415–428.

MacDonald, C., Lepage, K., Eden, U., & Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron*, *71*(4), 737-749.

Manning, J. R., Polyn, S. M., Baltuch, G. H., Litt, B., & Kahana, M. J. (2011). Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 12893-12897.

Marshall, L., Helgadottir, H., Mölle, M., & Born, J. (2006). Boosting slow oscillations during sleep potentiates memory. *Nature*, *444*, 610–613.

Marshall, L., Mölle, M., Hallschmid, M., & Born, J. (2004). Transcranial direct current stimulation during sleep improves declarative memory. *Journal of Neuroscience*, *24*, 9985-9992.

McFarlane, K. A., & Humphreys, M. S. (2012). Maintenance rehearsal: The key to the role attention plays in storage and forgetting. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*, 1001-1018.

McGaugh, J. L. (2000). Memory–A century of consolidation. *Science*, *287*, 248–251.

Mednick, S. C., Cai, D. J., Shuman, T., Anagnostaras, S., & Wixted, J. T. (2011). An opportunistic theory of cellular and systems consolidation. *Trends in Neurosciences*, *34*(10), 504-514.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*, 61-64.

Müller, G. E., & Pilzecker, A. (1900). Experimentelle Beiträge zur Lehre vom Gedächtnis [Experimental contributions to the science of memory]. *Zeitschrift für Psychologie*, *1*, 1–300.

Nairne, J., Neath, I., Serra, M., & Byun, E. (1997). Positional distinctiveness and the ratio rule in free recall. *Journal of Memory and Language*, *37*, 155-166.

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.

Oberauer, K. (2009). Design for a working memory. In B. H. Ross (Ed.), *The Psychology of Learning and Motivation, 51* (pp. 45–100). San Diego: Academic Press.

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, *55*, 601–626.

Oudiette, D., & Paller, K. A. (2013). Upgrading the sleeping brain with targeted memory reactivation. *Trends in Cognitive Sciences*, *17*, 142-149.

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

Raftery, A. E. (1996). Approximate Bayes factors and accounting for model uncertainty in generalised linear models. *Biometrica*, *83*(2), 251-266.

Rasch, B., Büchel, C., Gais, S., & Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science*, *315*, 1426-1429.

Sederberg, P. B., Gershman, S. J., Polyn, S. M., & Norman, K. A. (2011). Human memory reconsolidation can be explained using the temporal context model. *Psychonomic Bulletin and Review*. doi: DOI10.3758/s13423-011-0086-9

Shankar, K. H., & Howard, M. (2012). A scale-invariant internal representation of time. *Neural Computation*, *24*(1), 134-193.

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

Staresina, B., Alink, A., Kriegeskorte, N., & Henson, R. (2013). Awake reactivation predicts memory in humans. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(52), 21159-21164.

Tulving, E., & Arbuckle, T. (1966). Input and output interference in short-term associative memory. *Journal of Experimental Psychology*, *72*, 145–150.

Unsworth, N., Heitz, R. P., & Parks, N. A. (2008). The importance of temporal distinctiveness for forgetting over the short term. *Psychological Science*, *19*, 1078-1081.

Wixted, J. T. (2004). The psychology and neuroscience of forgetting. *Annual Review of Psychology*, *55*, 235–269.

Table 1

Constraints—lower and upper bounds—imposed on parameter estimation.

Parameter	Lower bound	Upper bound
c	1	100
t	0.1	1
s	1	100
o	0	1
w	0	1
α	0	100
b	0	3

Note. c, sharpness of memory representation; t, recall threshold; s, slope indicating noisiness of threshold; o, output interference; w, dimensional weight; α, consolidation growth factor, b, exponent determining nonlinear consolidation growth.

Table 2

Best-fitting models in Experiment 1.

Model	N(pars)	Deviance	$AICc$ wt	BIC wt
1d SIMPLE (no consolidation)	4	4569	.33	.38
2d SIMPLE (equally weighted dimensions, no consolidation)	4	4560	.36	.42
2d SIMPLE (free dimension weight, no consolidation)	5	4552	.12	.09
1d SIMPLE (linear consolidation)	5	4552	.13	.09
2d SIMPLE (free dimension weight, linear consolidation)	6	4548	.04	.02
2d SIMPLE (free dimension weight, non-linear consolidation)	7	4523	.02	.01

Note. 1d, one-dimensional; 2d, two-dimensional; N(pars), number of free model parameters; Deviance, summed deviance across all participants; $AICc$ and BIC wt, information criterion weights

Table 3

Model comparison and mean parameter estimates in Experiment 1.

Model	N(pars)	Deviance	$AICc$ wt	BIC wt	c	t	s	o	α
2d SIMPLE (equally weighted, no consolidation)	4	4560	.59	.70	21.26	.26	45.42	.64	-
1d SIMPLE (linear consolidation)	5	4552	.23	.17	44.82	.76	7.98	.63	.63
2d SIMPLE (equally weighted, no time, linear consolidation)	5	4595	.18	.13	30.02	.94	10.31	.45	.37

Note. 1d, one-dimensional; 2d, two-dimensional; N(pars), number of free model parameters; Deviance, summed deviance across all participants; $AICc$ and BIC wt, information criterion weights

Table 4

Best-fitting models in Experiment 2.

Model	N(pars)	Deviance	$AICc$ wt	BIC wt
1d SIMPLE (no consolidation)	4	4644	.32	.37
2d SIMPLE (equally weighted, no consolidation)	4	4633	.36	.42
2d SIMPLE (free weight, no consolidation)	5	4613	.14	.10
2d SIMPLE (equally weighted, linear consolidation)	5	4623	.13	.09
2d SIMPLE (free weight, linear consolidation)	6	4624	.04	.02
2d SIMPLE (free weight, non-linear consolidation)	7	4608	.01	$<.01$

Note. 1d, one-dimensional; 2d, two-dimensional; N(pars), number of free model parameters; $AICc$ and BIC wt, information criterion weights

Table 5

Model comparison in Experiment 2.

Model	N(pars)	Deviance	$AICc$ wt	BIC wt	c	t	s	o	α
2d SIMPLE (equally weighted, no consolidation)	4	4633	.63	.73	50.43	.27	34.93	.54	-
2d SIMPLE (equally weighted, linear consolidation)	5	4623	.23	.17	48.70	.30	32.00	.59	.23
2d SIMPLE (equally weighted, no time, linear consolidation)	5	4679	.14	.11	22.02	.89	8.70	.39	.25

Note. 1d, one-dimensional; 2d, two-dimensional; N(pars), number of free model parameters; $AICc$ and BIC wt, information criterion weights

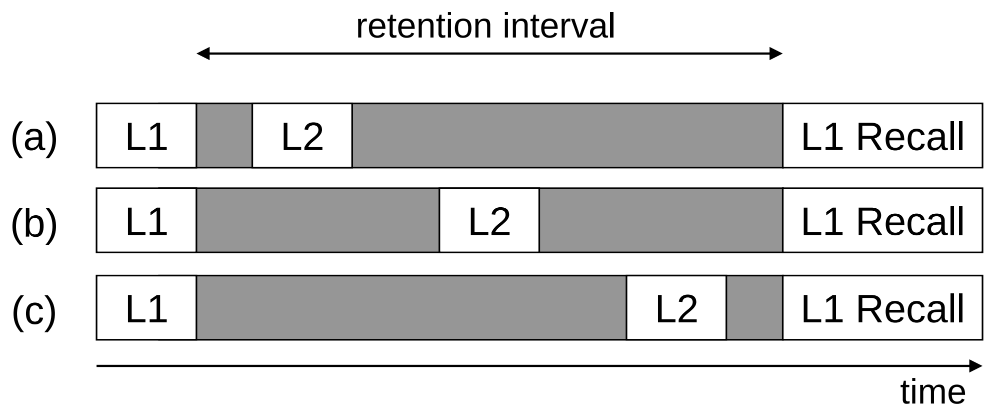

Figure 1. Schematic diagram showing various temporal positions of an interfering activity—study of list 2 (L2)—during the retention interval of to-be-remembered list 1 (L1).

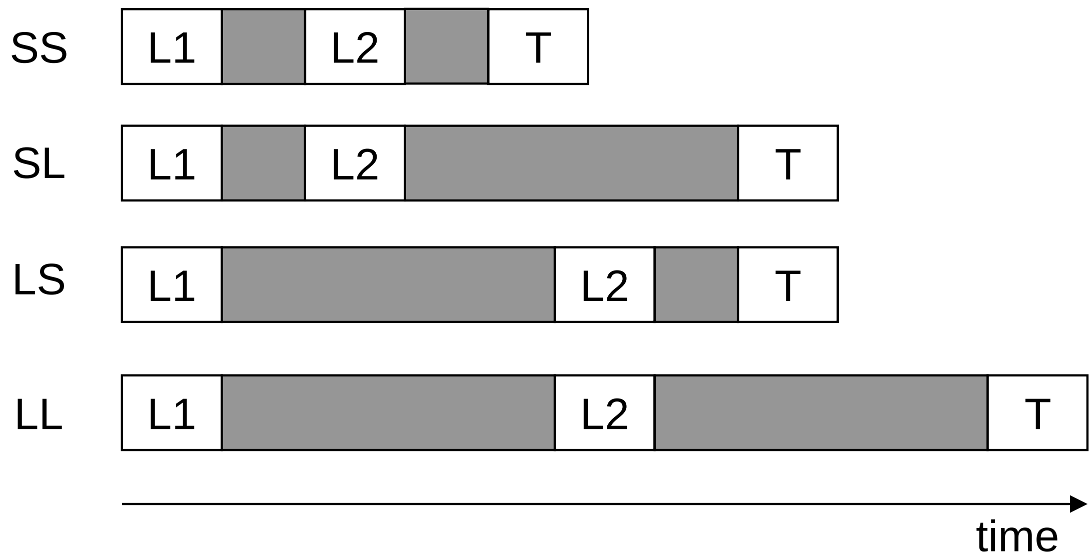

Figure 2. A schematic depiction of the four experimental condition used in both experiments. L1 and L2 denote the two study lists, which both involved 10 words, each presented at a rate of 1 per 900 ms. T denotes the recall test, which always targeted list 1; list 2 was subsequently recalled on a random half of the trials. Intervals between study/test events were filled with an easy tone-detection task (Experiment 1) or digit-verbalisation (Experiment 2) and lasted either 60 or 240 seconds. The four conditions are thus labeled SS (short L1-L2 interval, short L2-T interval), SL (short-long), LS (long-short), and LL (long-long). Participants completed four trials per condition.

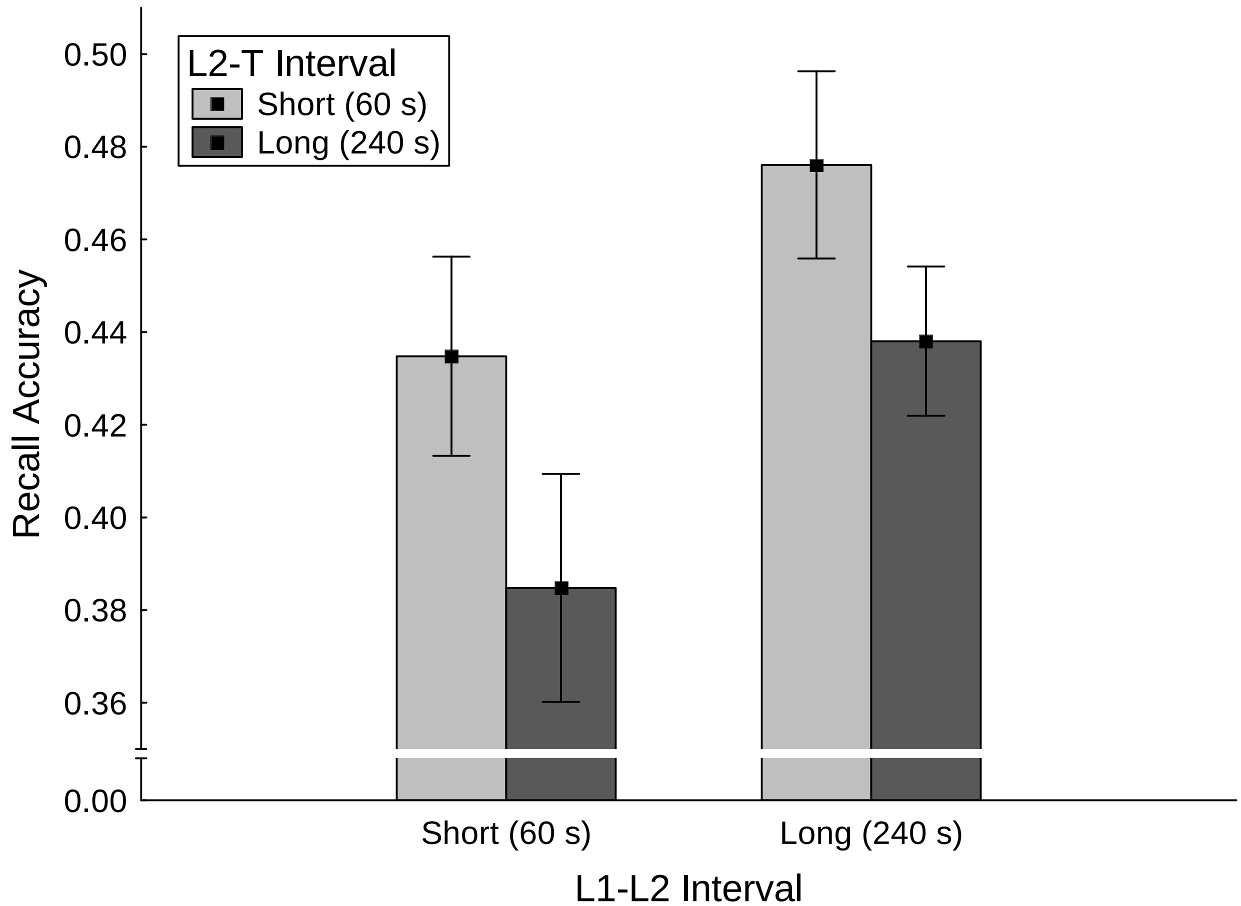

Figure 3. Mean recall accuracy across the four experimental conditions in Experiment 1. L1 and L2 refer to study lists 1 and 2, respectively. T refers to the recall test. Error bars depict standard errors of the mean (Morey, 2008).

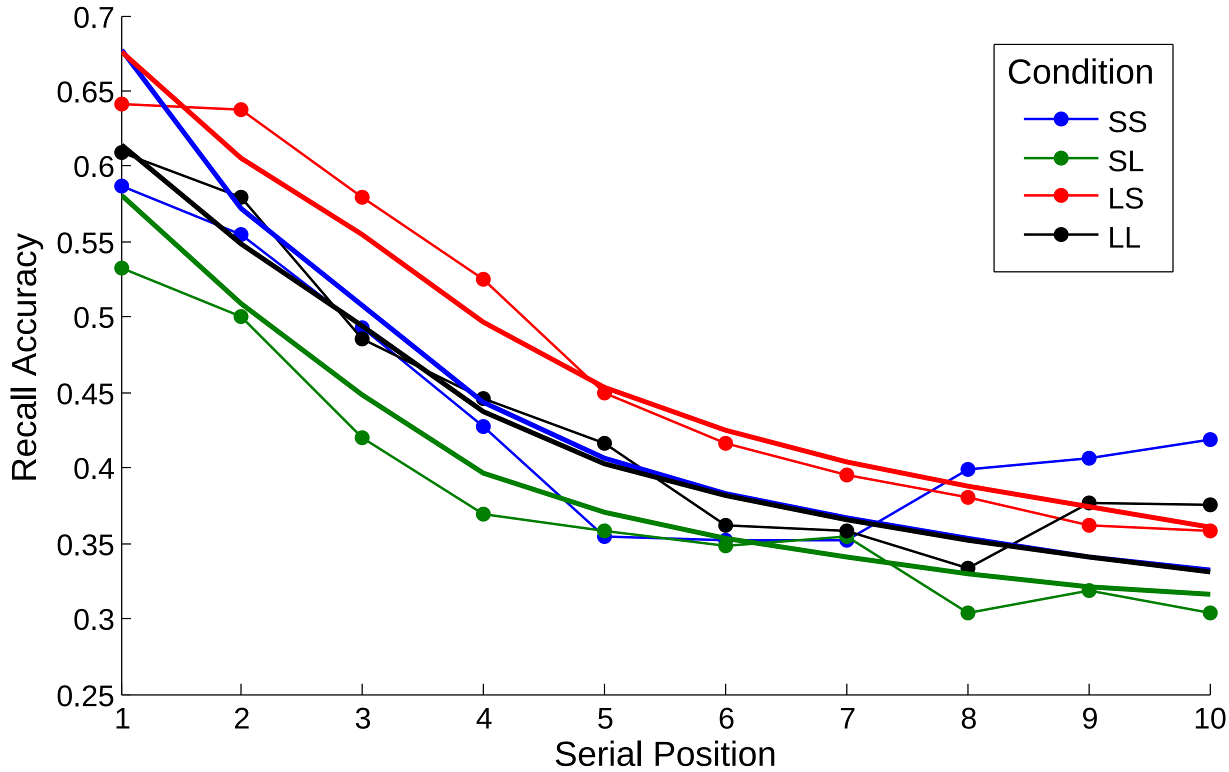

Figure 4. Smoothed serial position curves across the four experimental conditions in Experiment 1 (thin lines with data points) and the best-fitting model predictions (bold lines). Smoothing involved averaging the data of each serial position and its immediate neighbor(s). The four conditions are labeled SS (short L1-L2 interval, short L2-T interval), SL (short-long), LS (long-short), and LL (long-long).

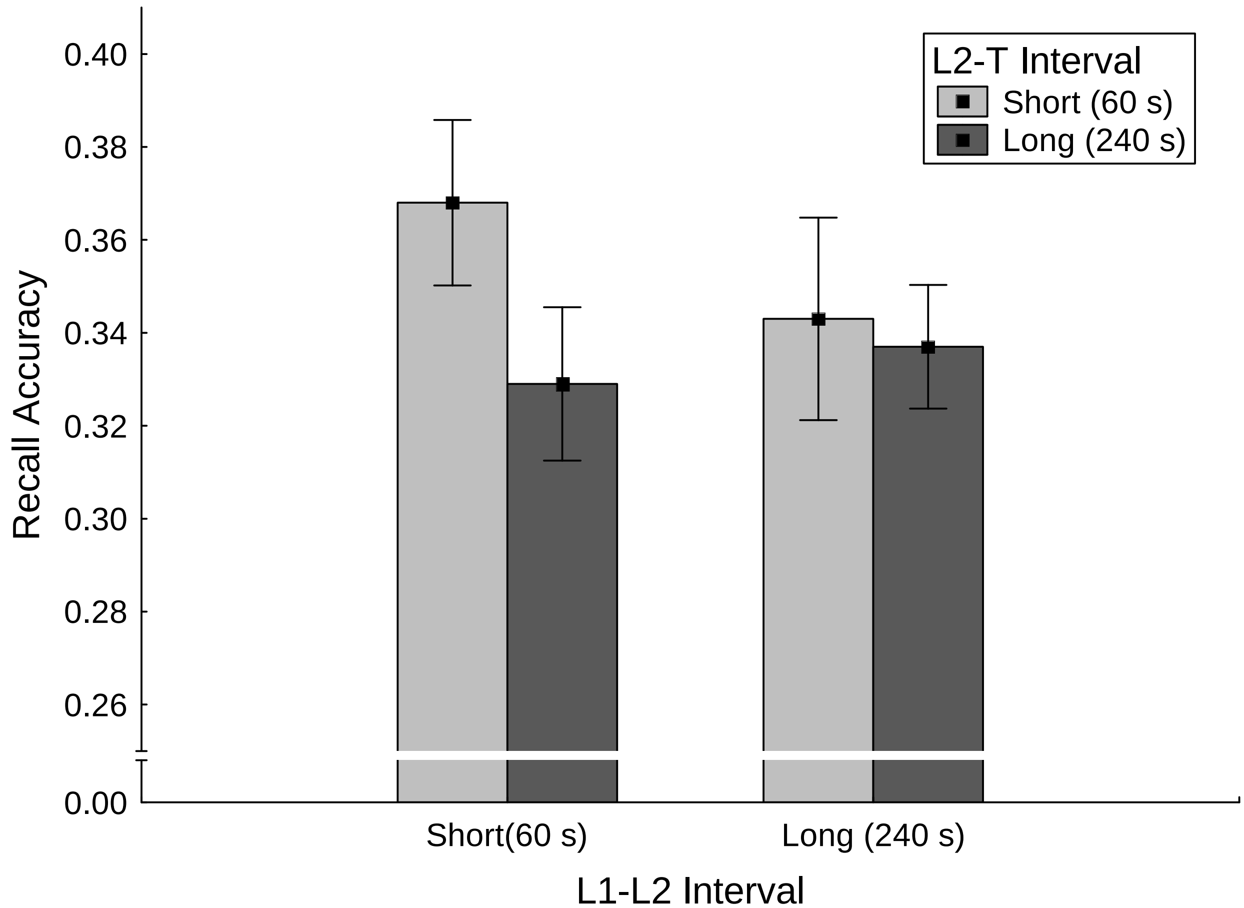

Figure 5. Mean recall accuracy across the four experimental conditions in Experiment 2. L1 and L2 refer to study lists 1 and 2, respectively. T refers to the recall test. Error bars depict standard errors of the mean (Morey, 2008).

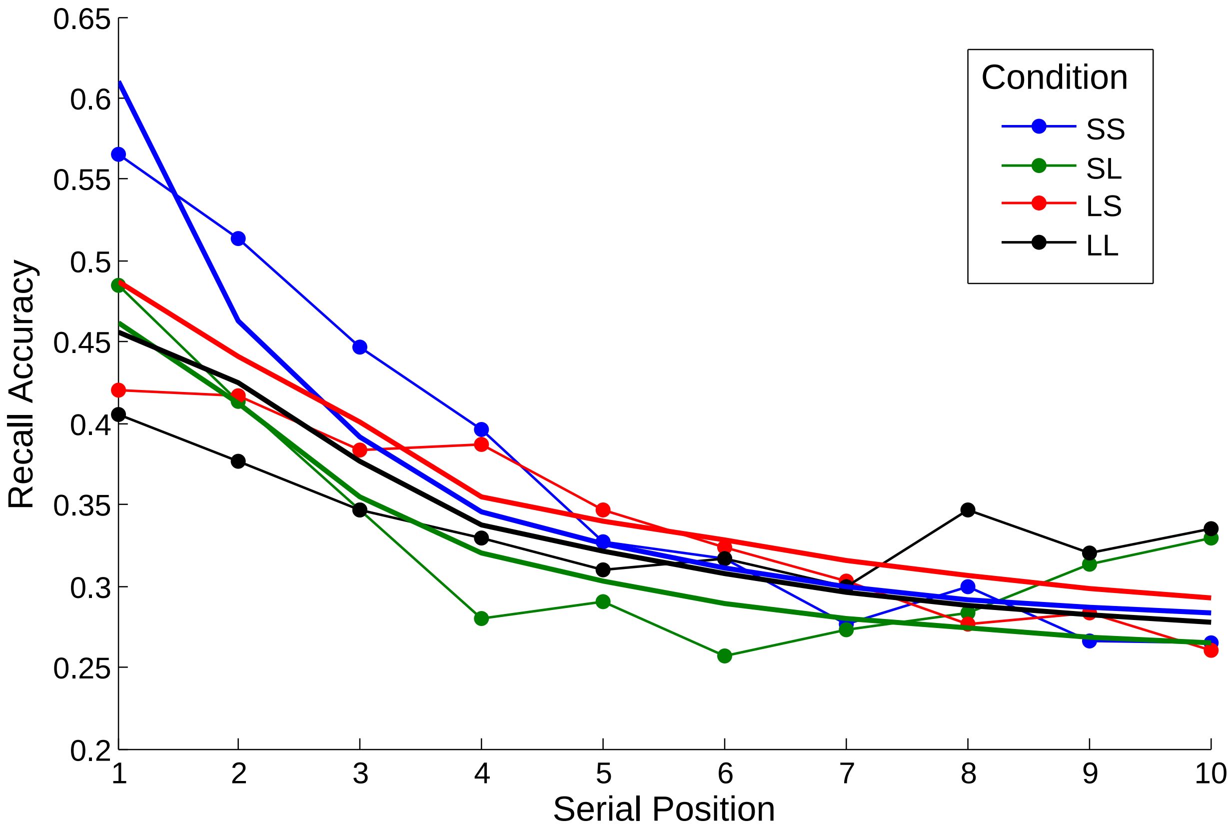

Figure 6. Smoothed serial position curves across the four experimental conditions in Experiment 2 (thin lines with data points) and the best-fitting model predictions (bold lines). Smoothing involved averaging the data of each serial position and its immediate neighbor(s). The four conditions are labeled SS (short L1-L2 interval, short L2-T interval), SL (short-long), LS (long-short), and LL (long-long).