
Volume 59, Number 2, August 2008
ISSN 0749-596X
ELSEVIER
Journal of
Memory and
Language
Editor
Kathryn Bock
Associate Editors
Robert Greene
Evan Heit
James McQueen
Debra J. Jared
Editorial Board
Gerry T.M. Altmann
Sally Andrews
Julie E. Boland
Gordon Brown
Jenn-Yeu Chen
Charles Clifton, Jr.
Cynthia M. Connine
Gary S. Dell
Kara D. Federmeier
Fernanda Ferreira
Victor Ferreira
Cynthia Fisher
Kenneth I. Forster
Carol A. Fowler
David A. Gallo
Alan Garnham
Simon Garrod
Richard J. Gerrig
Zenzi M. Griffin
Robert J. Hartsuiker
Larry L. Jacoby
Colleen Kelley
Judith F. Kroll
Wido La Heij
Stephen J. Lupker
Michael E.J. Masson
Brian McElree
Antje S. Meyer
Akira Miyake
Gregory L. Murphy
James S. Nairne
James H. Neely
Martin J. Pickering
Suparna Rajaram
Keith Rayner
Henry L. Roediger III
Arthur G. Samuel
Patrick Sturt
Daniel Swingley
Marcus Taft
Rebecca Treiman
John Trueswell
Jos J.A. van Berkum
Mieke Verfaillie
Andrew Yonelinas
Available online at
ScienceDirect
www.sciencedirect.com

Journal of Memory and Language 59 (2008) 200–222

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Interference-based forgetting in verbal short-term memory

Stephan Lewandowsky [a,*], Sonja M. Geiger [a], Klaus Oberauer [b]

[a] School of Psychology, University of Western Australia, Crawley, WA 6009, Australia
[b] Department of Experimental Psychology, 12a Priory Road, Bristol, BS8 ITU, UK

ARTICLE INFO

Article history:
Received 20 January 2008
revision received 10 April 2008
Available online 6 June 2008

Keywords:
Short-term memory
Computational models of memory
Forgetting

ABSTRACT

This article presents four experiments that tested predictions of SOB (Serial Order in a Box), an interference-based theory of short-term memory. Central to SOB is the concept of novelty-sensitive encoding, which holds that items are encoded to the extent that they differ from already-encoded information. On the additional assumption that distractors are encoded into memory in the same manner as list items, the theory predicts differential effects of interfering activity based on the similarity structure of distractors. Consistent with predictions, three experiments showed that overt articulation of distractors in between recalls of list items did not affect forgetting when the same distractor was repeated multiple times, whereas forgetting was observed if several different distractors were articulated within the same time span. A fourth experiment showed that the absence of forgetting with repeated articulations of the same item was not due to compensatory attentional refreshing of memory traces. The data support the notion that forgetting from short-term memory arises from interference and are difficult to reconcile with temporal decay.

There has been much renewed interest in the causes of forgetting from verbal short-term memory, in particular whether forgetting arises from temporal decay or event-based interference. Some theorists have proposed that memory representations inexorably and rapidly fade over time (e.g., Page & Norris, 1998), whereas others have ascribed forgetting to the interference that arises from the processing of subsequent events (e.g., Farrell & Lewandowsky, 2002; Lewandowsky & Farrell, in press). The latter view acknowledges that forgetting is necessarily *expressed* over time, but it denies that time plays a *causal* role in forgetting.

Several sources of evidence have been explored to adjudicate between these alternative views. For example, many decay theorists present the "word-length effect" as a crucial supporting piece of evidence (e.g., Cowan, 1995). The word-length effect was first reported by Baddeley, Thomson, and Buchanan (1975) and refers to the now well-established finding that lists composed of long words (e.g., "hippopotamus", "establishment") are recalled less accurately than lists of short words (e.g., "cut", "hip"). As we have argued in detail elsewhere (Lewandowsky & Oberauer, in press), notwithstanding its intuitive appeal, the word-length effect does not provide evidence for time-based decay. Our argument there was twofold: first, the word-length effect is correlational and hence necessarily confounded with a number of other variables that give rise to alternative explanations. Second, an analysis of potential non-verbal modes of memory refreshing raises doubts as to whether a decay model would necessarily predict the word-length effect in the first place.

Forgetting and the distractor methodology

An alternative source of evidence concerning the causes of forgetting, involving the direct experimental manipulation of retention time, may therefore present a more promising avenue (e.g., Cowan et al., 2006; Cowan, Nugent,

* Corresponding author.
E-mail addresses: lewan@psy.uwa.edu.au (S. Lewandowsky), K.Oberauer@bristol.ac.uk (K. Oberauer).
URL: http://www.cogsciwa.com (S. Lewandowsky).

0749-596X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jml.2008.04.004

Elliott, & Geer, 2000; Cowan, Wood, Nugent, & Treisman, 1997; Lewandowsky, Duncan, & Brown, 2004). Although it is in principle straightforward to manipulate time, for example by simply pacing the rate at which items are presented (Lewandowsky, Brown, Wright, & Nimmo, 2006; Nimmo & Lewandowsky, 2005), temporal manipulations open up the possibility of compensatory rehearsal. There is considerable agreement in the literature that articulatory rehearsal can neutralize the effects of decay (e.g., Baddeley, 1986), and hence any unfilled post-presentation intervals may entail rehearsal of already-presented material, thus masking whatever decay might otherwise have occurred during that time.

A potential resolution to this methodological problem was presented by Lewandowsky et al. (2004). In their studies, participants recalled a short list of letters in forward order while repeating a single irrelevant word aloud in between retrievals. There is broad consensus that this "articulatory suppression" (AS from here on) blocks rehearsal (Baddeley, 1986, p. 37 & 86; Baddeley & Lewis, 1984; Page & Norris, 1998, p. 764 & 770), thus providing an opportunity for decay to express itself if it exists. Contrary to the expectations of the decay view, Lewandowsky et al. found that recall performance was unaffected by the number of times the irrelevant word was repeated in between memory retrievals, notwithstanding the considerable increase in retention time with increasing numbers of repetitions of the irrelevant word. By implication, time per se did not cause forgetting in this instance.

The results of Lewandowsky et al. were replicated and extended by Oberauer and Lewandowsky (in press), who compared a distractor-free baseline condition to a variety of conditions with filled delays during retrieval. (Their studies also included an encoding manipulation which is not relevant here.) In some conditions, AS in between retrievals was additionally accompanied by a speeded choice task (in which two typographic symbols had to be identified via arbitrarily-mapped response keys) to prevent not only articulatory rehearsal but also a recently-identified second rehearsal mechanism, known as attentional refreshing (Hudjetz & Oberauer, 2007; Raye, Johnson, Mitchell, Greene, & Johnson, 2007). The data showed that the duration of the interfering activity had at most a negligible effect on performance, even when AS was accompanied by the speeded choice task. Oberauer and Lewandowsky applied several computational models to their results and found that two alternative decay models (the Primacy Model; Page & Norris, 1998; and a model based on association of items to positional markers) could not handle the data. By contrast, an interference-based model (SOB; Farrell & Lewandowsky, 2002; Lewandowsky & Farrell, in press) gave a better quantitative account of the data. Specifically, both decay models predicted substantial forgetting with increasing numbers of distractors after each retrieval attempt. Because forgetting cumulates over retrieval attempts, it affects items increasingly toward the end of the serial position curve. In consequence, decay models predict that the serial position curves should fan out, with a steeper decline when more distractors follow each retrieval. In contrast, SOB correctly predicted that the number of distractors has at most a small effect, thus leading to little difference between the serial position curves of trials with one and trials with several distractors following each item.

Although the results and modelling just reviewed point away from temporal decay as the sole source of forgetting in verbal short-term memory, the data have also created a new empirical and theoretical puzzle. In particular, how can the results of Lewandowsky et al. (2004) and Oberauer and Lewandowsky (in press) be reconciled with the plethora of evidence (beginning with Peterson & Peterson, 1959) that memory traces are increasingly forgotten as the retention interval, filled with some distracting activity, is increased?[1] And how can the results of Lewandowsky et al. (2004) be reconciled with reports of increasing forgetting as the number of distractors was increased even within the very same experimental paradigm (Page, 2006)? In a nutshell, the available data sometimes show forgetting over a filled retention interval and sometimes they do not, and the two outcomes can occur within seemingly similar circumstances within the same paradigm. By implication, a complete account of the existing data requires a theory that can turn forgetting on and off based on a principled mapping with the experimental circumstances.

This article presents simulations and four experiments that reconcile the seemingly puzzling presence and absence of forgetting in verbal short-term memory within a common empirical and theoretical framework. By way of brief overview, we begin by presenting quantitative predictions of the SOB model that were derived from its successful application to the results of Oberauer and Lewandowsky (in press). SOB predicts that whether or not forgetting is observed during a distractor interval depends crucially on the type of distracting material: if the same distractor is repeated multiple times, as in the studies of Lewandowsky et al. (2004) and Oberauer and Lewandowsky (in press), additional repetitions of that distractor should not lead to (much) further forgetting. If the identity of the distractor changes within a set of repetitions in between retrievals, the model expects memory to decline with the number of distractors under otherwise identical circumstances. We then report three experiments whose consistent results are in accord with SOB's predictions: increased forgetting over longer filled retention intervals was observed whenever different words had to be articulated as distractors, whereas lengthening of filled intervals had no effect on memory under otherwise identical circumstances when the same distractor was repeated multiple times. A fourth experiment shows that this difference did not reflect the comparable ease with which the same item can be repeatedly spoken, because repetitions of the same word failed to yield forgetting even if they were accompa-

[1] At first glance, one might be tempted to link the occurrence of forgetting to the experimental paradigm: the Brown–Peterson paradigm is widely known to yield rapid and pervasive forgetting, suggesting perhaps that filled retention intervals always engender forgetting. By implication, one might conclude that there is something peculiar about interspersing distractors in between retrieval attempts that prevents the emergence of forgetting. As we show in the remainder of this article, this possibility is incorrect: forgetting can readily be obtained with interspersed distractors, and conversely, there is precedent for the *absence* of forgetting in the Brown–Peterson paradigm.

nied by a demanding additional task. We conclude that forgetting in verbal short-term memory is due to interference rather than decay.

Novelty-sensitive encoding and forgetting in SOB

Farrell and Lewandowsky (2002) presented a distributed model of short-term recall that they called SOB, for "Serial-Order in a Box", a name that acknowledged the model's reliance on the Brain-State-in-a-Box algorithm (e.g., Anderson, Silverstein, Ritz, & Jones, 1977). Briefly, SOB assumes that items are represented by vectors of features that are encoded into memory by adding their auto-associations to a common weight matrix. All encoding and retrieval processes in SOB are entirely dependent on events only and not affected by the passage of time. Thus, whatever forgetting occurs in SOB results from interference not temporal decay.

The basic architecture of the model involves the continual super-imposition of new information onto items already presented. Most relevant here is the core property of SOB that encoding strengths are a function of the novelty of incoming items. Every time an item is presented for study, its novelty is first assessed by comparison to the weight matrix.[2] If the item is judged to be novel, it is then encoded with a large encoding weight, whereas if it resembles already-encoded information, its encoding weight is considerably smaller. In the extreme case of an exact repetition, an item's encoding weight is negligible.

This encoding mechanism, which has been a crucial part of the theory since its inception, has several important consequences. For example, SOB automatically creates a primacy gradient across serial positions because the first list item is necessarily more novel than all subsequent ones (Farrell & Lewandowsky, 2002). Novelty-sensitive encoding also predicts a mixed-list advantage for dissimilar items embedded in a list of phonologically similar items (Lewandowsky & Farrell, 2008), a prediction that has been confirmed repeatedly (Farrell, 2006; Farrell & Lewandowsky, 2003; Lewandowsky & Farrell, 2008). Finally, and most crucially in the present context, novelty-sensitive encoding has also enabled SOB to handle the absence of forgetting that has hitherto been associated with distractor activity during retrieval.

Oberauer and Lewandowsky (in press) applied SOB to their data (and by implication to those of Lewandowsky et al., 2004) by assuming that all distractor items were encoded into memory during recall in much the same way as list items were encoded during study. Mirroring the way in which experimental participants articulated a distractor either once or multiple times, the representation of the distractor word was added either once or several times in between retrievals. Owing to novelty-sensitive encoding, these assumptions captured the known effects of distractors at retrieval: the first time a distractor is encountered, it receives a large encoding weight because SOB identifies it as novel. In consequence, the large update to the weight matrix is necessarily disruptive of existing memories. This captures the fact that compared to a no-distractor baseline list, a single distractor causes a dramatic performance loss (Oberauer & Lewandowsky, in press). By the same token, if there are additional repetitions of the same distractor, those repetitions receive minimal encoding because they are identified as having virtually no novelty. In consequence, there is little additional interference compared to the single-distractor condition, exactly as observed in the data. Because distractor(s), like list items, are associated to a context marker that evolves across output positions, the first occurrence of a distractor after the next recall is again considered novel and hence interfering.

The novelty-sensitive encoding of distractors immediately implies strong—and as yet untested—predictions involving the similarity structure of the to-be-articulated distractor material. Consider an extreme case in which every distractor consists of a different word; that is, people might say "table–horse–truck" after recall of the first item and "orange–zucchini–car" after the second item and so on. Within the novelty-sensitive encoding framework, every distractor in this case is relatively novel and would be encoded quite strongly, hence maximizing the amount of interference and giving rise to forgetting in proportion to the number of distractor events. It follows that the existing successful application of SOB to the distractor-based forgetting methodology entails a principled built-in linkage between situations in which forgetting does and does not occur, without parameter change and based exclusively on analysis of the distracting material. In a nutshell, forgetting should be absent with repeated articulations of the same word and it should emerge over an identical time scale with the articulation of a set of different words.

We obtained quantitative predictions along those lines from SOB using the published parameter settings of Oberauer and Lewandowsky (in press, Experiment 2). The only changes made to the simulation related to the methodology for the present experiments; list length was reduced from 6 to 5 and the first burst of distractors occurred after retrieval of the first item rather than before (We use the term "burst" from here on to refer to the variable number of distractors that are articulated in between retrievals.). Further details about the simulation and SOB, including a description of how distractors were generated, are provided in the Appendix.

We derived predictions for four distinct similarity structures among distractors by orthogonally combining the extent of *between-burst* similarity (i.e., whether or not the same distractor(s) were articulated in between retrievals) with the extent of *within-burst* similarity (i.e., whether each burst involved the same or different items). The second similarity variable becomes relevant only when bursts contain more than one distractor.

The predictions of SOB are summarized in Fig. 1, with rows of panels corresponding to different levels of within-burst similarity and columns referring to different between-burst similarities. The three serial position curves in each panel represent a distractor-free baseline condition and lists with bursts of 1 and 3 distractors, respectively. The top panels show predictions for bursts involving repe-

[2] The novelty of an item is measured by its "energy" with respect to the weight matrix; it is for this reason that the encoding mechanism in SOB is also called "energy-gated". Here, we prefer the less technical term "novelty-sensitive" encoding.

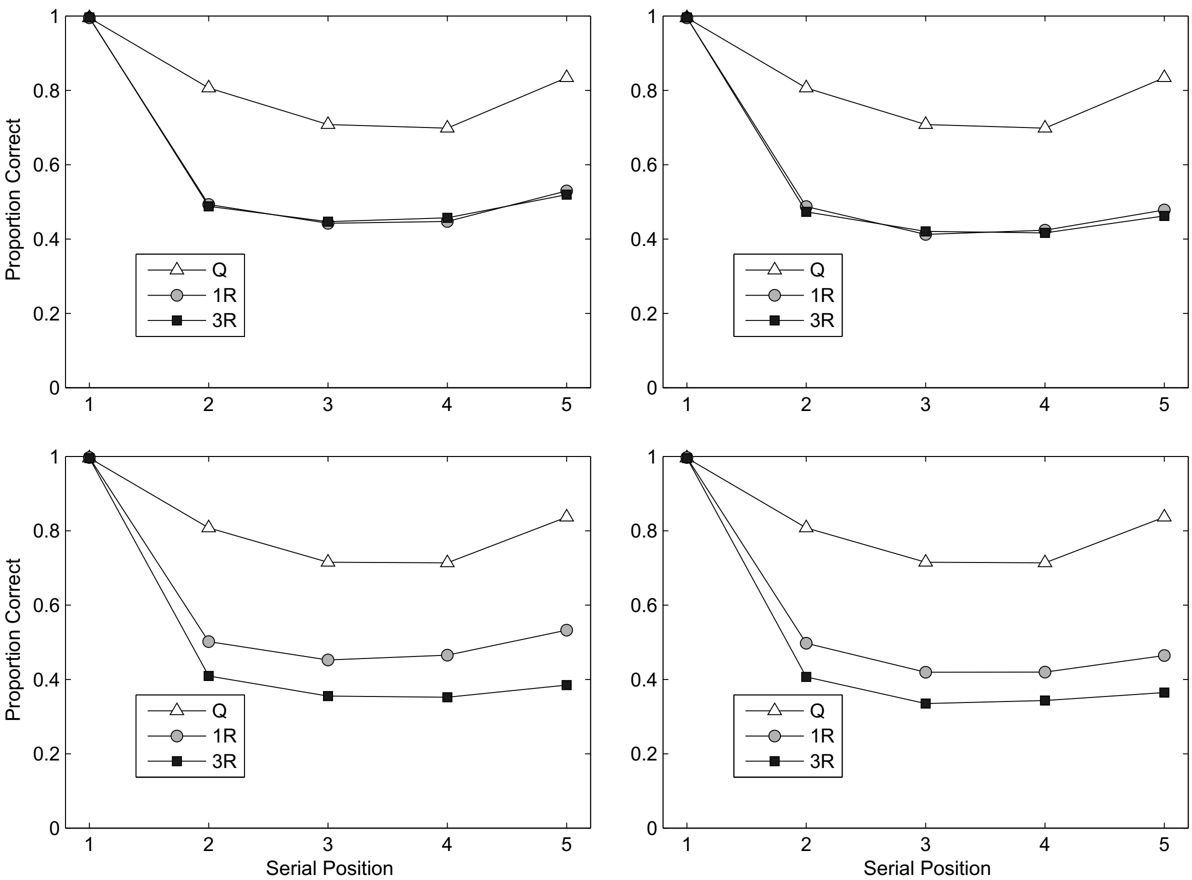

Fig. 1. Serial position curves predicted by SOB for the distractor manipulations explored in this article (Q refers to distractor-free quiet baseline condition, 1R and 3R refer to 1 and 3 distractors, respectively, in between each pair of retrieval; the same labeling is used in all figures). The upper panels show predictions for simple bursts and the lower panels for the complex bursts. Steady bursts are shown in the left-hand panels and changing bursts in the right-hand panels; see text for more explanation. Details of the simulation are provided in the Appendix.

tition of the same word, whereas the bottom panels show predictions for bursts consisting of three different words (in the case of three-word bursts). From here on, we refer to the former type of burst as *simple* (e.g., "super…super …super") and the latter (e.g., "quokka…marron…bilby") as *complex*.

The left-hand panels show the predictions when the same bursts—irrespective of whether they are simple or complex—are repeated across all output positions. We refer to this level of between-burst similarity as *steady* from here on. The right-hand panels, by contrast, show predictions when the identity of bursts changed across output positions, a level of between-burst similarity that we call *changing* from here on. Note that changing between-burst similarity can involve either simple bursts (in which case "super…super…super" might be followed by "quokka…quokka…quokka" after the next recalled item) or complex bursts ("quokka…marron…bilby" followed by "alligator…hemlock…garden").

To summarize, the top-left panel corresponds to people saying the same word either once or three times in between all retrievals; the top-right panel represents people saying a word either once or three times, with the identity of that word changing across output positions. Finally, the bottom panels correspond to people articulating three different words at each output position, with the identity of those three words either repeating across output positions (left-hand panel) or changing continuously (right-hand panel).

The predictions of SOB can be summarized as follows: (1) within-burst similarity has a substantial effect on the amount of forgetting over filled delays. For simple bursts, the number of distractors is not expected to have any effect: articulation of the same item three times in between retrievals should not impair memory any further than saying that word once, notwithstanding the additional time that would necessarily elapse in between retrievals. Averaging across serial positions, SOB predicts a minute reduction in performance of less than 1% for both steady (.581 vs. .581 for 1 and 3 distractors, respectively) and changing (.559 vs. .554) simple bursts.

For complex bursts, by contrast, SOB predicts a substantial effect of the number of distractors that can be broken into two distinct components. (a) SOB predicts a notable main effect across serial positions (.59 vs. .50 for steady and .56 vs. .49 for changing complex bursts). (b) The effect of the number of distractors is modulated by serial position and involves a "fanning out" from a common origin. To clarify the functional form of that fanning, Fig. 2 shows the amount of predicted forgetting expressed as the difference in performance with 1 vs. 3 distractors. The fanning of serial position curves that is apparent in Fig. 1 for complex

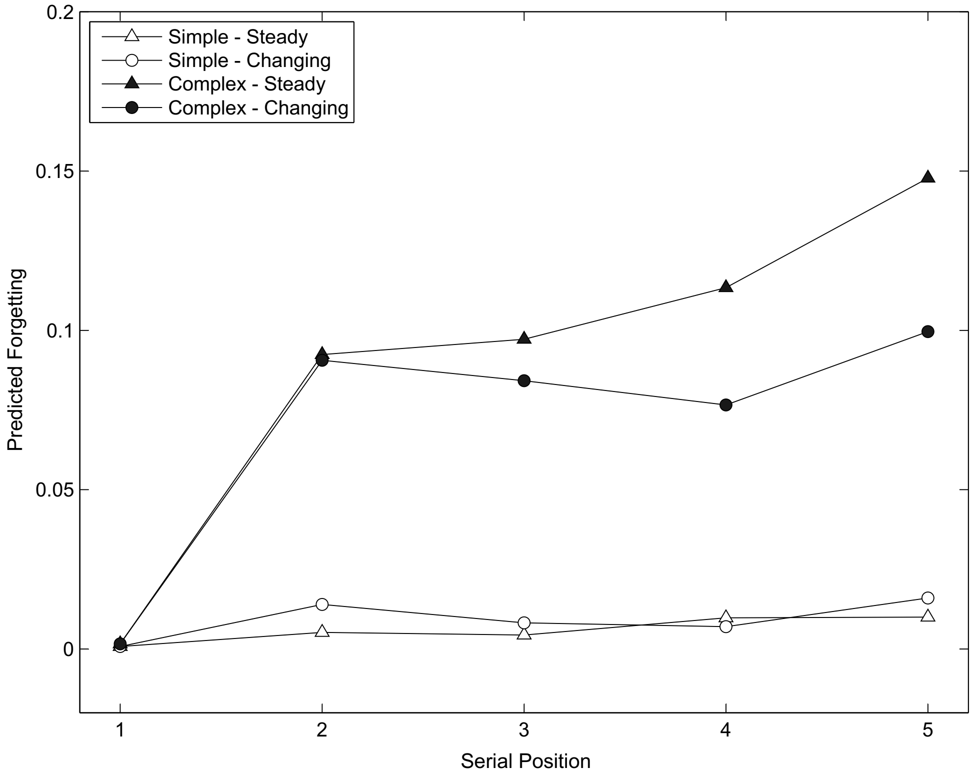

Fig. 2. Predictions of SOB shown as a function of the difference in performance between lists with 1 and 3 distractors (i.e., showing net distractor-induced forgetting) across serial positions for each of the four conditions in Fig. 1.

bursts is represented by the increase in predicted forgetting across serial position in Fig. 2. The increase in forgetting is negatively accelerated and, in the case of changing complex bursts, even expected to decrease slightly beyond the second serial position.

(2) SOB also predicts a very small main effect of between-burst similarity. When the identity of a single distractor changes across output positions, performance is impaired slightly more than when the same item is used at each position. Numerically, the average difference between the steady and changing serial position curves for a single distractor is .033 across serial positions 2 through 5 (This difference is the same irrespective of whether bursts are simple or complex because the two are indistinguishable for a single distractor.).

(3) Within-burst similarity and between-burst similarity interact under-additively in their effect on forgetting over filled delays. That is, the extent of forgetting produced by complex bursts is greater if bursts are repeated across output positions than when they are not (compare the fanning in the bottom two panels). To illustrate the rather small magnitude of the effect, the difference between the serial position curves for 1 and 3 complex distractors is .113 (averaged across serial positions 2 through 5) with steady bursts, but .088 for changing bursts. Note that this effect arises largely because steady and changing bursts differ for a single distractor, rather than because of differences for bursts of three distractors.

These predictions arise directly from the core principles of SOB: Identical distractors within a burst are encoded with decreasing strength due to novelty-sensitive encoding, whereas dissimilar distractors receive full encoding strength. Therefore, complex bursts cause a larger weight change in the association matrix, thus effectively reducing the signal strength of studied items. This effect is attenuated for early serial positions because those items were themselves stored more strongly owing to novelty-sensitive encoding and are thus more resistant to interference. The effect of between-burst similarity arises in the same way. A distractor repeated in a different output position is still recognized as less novel than a new distractor, and its encoding is therefore dampened. Between-burst similarity has a smaller effect than within-burst similarity, because distractors at different serial positions are associated to different positional markers, and as a consequence, distractors repeated in different bursts at different output positions are regarded as more novel than distractors repeated within the same burst. We now present four experiments that explored these predictions of SOB.

Experiment 1

The first experiment used a predictable and well-rehearsed sequence of distractors; namely, the months of the year. This material was chosen to minimize the contribution of potentially differing demands on general cognitive resources to the effects of within- and between-burst similarity (Experiment 4 below provides additional empirical assurance that cognitive resources play no contaminating role in the first three studies).

The predictability of distractors has been shown to affect memory performance at least in children (Gavens & Barrouillet, 2004). The predictions of SOB derive from the similarity structure of distractors, not their predictability, and therefore we made an effort to equate predictability between conditions. People studied lists of 5 letters for immediate forward serial recall. During recall, overt articulation of 1 or 3 distractors (i.e., names of months) was inserted in between oral reports of list items. The within-burst similarity (simple vs. complex bursts) was manipulated between subjects, whereas the remaining two variables (number of distractors, 1 vs. 3; between-burst similarity, whether the identity of distractors changed or were steady across output positions) were manipulated within subjects.

Methods

Participants

Participants were 24 members from the University of Western Australia campus community. They participated voluntarily in a 1-h session for remuneration of A$10. An equal number of participants were randomly assigned to each of the two experimental conditions.

Materials and procedure

Participants had to study and recall lists of five consonants in correct order. List items were sampled randomly without replacement from a set of 19 letters (all consonants except Q and Y). A Windows computer running a Matlab program, designed using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997), was used to display stimuli and record responses for all studies reported here.

Each trial commenced with presentation of a fixation cross for 0.9 s in the center of the screen, which was followed by sequential presentation of the list items at the rate of 1 letter/s (900 ms on, 100 ms off). Items were displayed centrally in black on a white background. The last list item was immediately followed by a cue that specified the number and nature of distractors, if any, on that particular trial (e.g., "3 changing" or "1 steady"). The cue was visible for 1.5 s and was followed by a question mark which prompted oral report of the first list item.

People articulated either 0, 1, or 3 distractors after each retrieval and the identity of distractors either changed or remained the same across output positions, yielding five different list types: a baseline condition with no distractors, plus the orthogonal combination of 1 or 3 distractors that were either changing or steady across serial positions. Participants in the complex condition articulated three consecutive months (e.g., "January", "February", "March") whenever three distractors were required, whereas participants in the simple condition repeated the same month three times. Irrespective of whether bursts were simple or complex, each consecutive output position involved consecutive (sets of) months when distractors were changing, or the same (set of) month(s) when distractors were steady. Thus, in the complex-burst group the list type "3 changing" would involve, for instance, the distractors "January", "February", "March" after the first retrieval attempt, followed by "April", "May", "June" after the second retrieval event, and so on. In the simple-burst group this list type would involve the distractors "January", "January", "January" after the first retrieval attempt, followed by "February", "February", "February" after the second.

Each of the 12 months served as a starting month for one randomly chosen participant within each condition. The starting month remained the same across all trials for a given participant. If required (e.g., if December was the starting month), the sequence of distractors "wrapped around" (i.e., December was followed by January, etc.).

During recall, each oral report of an item was recorded by the experimenter on the keyboard and was followed 200 ms later by test of the next output position. If a list involved distractors, each report after the first one was preceded by rapid presentation (200 ms per distractor) of the distractor(s) at the top of the screen (printed black on white, with each distractor on a separate row). Distractors were followed by presentation of the '?' prompt, which remained visible until the experimenter entered the next recalled list item. Participants read the distractors aloud as quickly as possible and without pausing before recalling the next letter as quickly as possible.

The experimental trials were preceded by six practice lists which were shown in increasing order of difficulty (baseline followed by steady-distractor and then changing-distractor trials). The sequence of 60 experimental trials (12 of each list type) was randomized anew for each participant. A self-paced break was administered after every 15 experimental trials.

Results

Individual differences

Correct-in-position recall performance was first examined at the level of individual participants aggregated across all within-subjects conditions. Individual performance ranged from .55 to .91, safely away from guessing level and from ceiling, and thus all participants were retained for analysis.

Timing manipulation

To test the success of our timing manipulation, we first considered the mean response times for recalling a letter, measured as the interval from presentation of the first distractor word (or question mark in the baseline condition) to the experimenter's key press for output positions 2–5. (The first recall was not preceded by distractors in any of the conditions and does not meaningfully contribute to this analysis, hence that output position was not considered here.)

Response times for all experiments are summarized in Table 1, averaged across output positions 2 through 5 and separated by condition. Compared to baseline, speaking a single distractor aloud took approximately one additional second, and speaking two more distractors increased response times by one more second irrespective of within-burst similarity. It must be noted that those temporal differences cumulate across serial positions; hence the 1-s difference between 1 and 3 distractors at each output position translates into an additional delay of 4 s by the time the entire list has been recalled. Given that decay theorists have equated memory span with the amount of verbal material that can be articulated in about 2 s (Schweickert & Boruff, 1986), our manipulation therefore proved sufficient to evaluate the decay hypothesis.

Table 1
Response times (s) averaged across output positions 2–5 in all four experiments

Distractors	Baseline	1 steady	1 changing	3 steady	3 changing
Experiment 1					
Simple	1.23	2.18	2.25	3.08	3.13
Complex	1.10	2.21	2.34	3.08	3.28
Experiment 2					
Simple	1.01	1.97	2.15	2.79	2.87
Complex	1.09	2.23	2.38	3.03	3.42
Experiment 3					
Familiar	1.10	2.03	2.18	2.91	3.11
Unfamiliar	1.15	2.13	2.27	3.17	3.48
Experiment 4					
Simple + Choice task		2.46		5.09[a]	

[a] In Experiment 4 there were 4 rather than distractors.

Response times in the baseline condition differed significantly from those with one distractor in each group; $F(1,11) = 118.5$, $p < .0001$, partial $\eta^2 = .92$, for the simple-burst group and $F(1,11) = 217.8$, $p < .0001$, partial $\eta^2 = .95$, for the complex bursts (contrasts conducted within a one-way within-subjects ANOVA involving all five conditions). Likewise, the comparison between one and three distractors was significant in each group, $F(1,11) = 118.9$, $p < .001$, partial $\eta^2 = .92$ for the simple-burst group, and $F(1,11) = 268.2$, $p < .001$, partial $\eta^2 = .96$, for the complex-burst group (using a $2 \times 2 \times 4$ ANOVA without the baseline but including between-burst similarity, number of distractors, and serial position as factors). In the complex-burst group alone, response times were somewhat longer for changing than for steady distractors, $F(1,11) = 8.5$, $p = .01$, partial $\eta^2 = .44$. We conclude that varying the number of distractors between recall attempts successfully and nearly equally delayed responding for both simple and complex bursts.

Recall performance. All analyses were based on correct-in-position scoring. To compare the general performance level of the two groups, a 2 (within-burst similarity group: simple vs. complex) $\times$ 5 (serial position) between-within ANOVA including only the baseline lists was conducted, which revealed neither a main effect of group, $F(1,22) = 1.17$, partial $\eta^2 = .05$, nor an interaction of group with serial position, $F(4,88) < 1$, partial $\eta^2 = .02$.

Fig. 3 shows the serial position curves for both groups and all list types using the same layout as for the SOB predictions. All error bars in figures in this article represent standard errors computed using the within-subjects procedure proposed by Bakeman and McArthur (1996) and are therefore only indicative of within-subjects comparisons (e.g., all those within a panel). For ease of exposition, the effects of our distractor manipulations were analyzed with two separate within-subjects ANOVAs for the simple- and the complex-burst group, respectively, with serial position (1–5), number of distractors (1 vs. 3), and between-burst similarity (steady vs. changing) as independent variables. Note that the baseline lists were omitted from these analyses. Note also that all conclusions reported in this article are supported by the appropriate joint analysis of both conditions in an omnibus ANOVA; we selectively report the crucial interactions from the omnibus analyses below.

Results of both analyses of variance are summarized in Table 2. Our main interest is in whether serial position curves for 1 versus 3 distractors fan out in the way characteristic of increased forgetting with a longer filled delay. The nature of the predictions provided by SOB suggested that this would be statistically detectable in two ways: first, by a main effect of the number of distractors for complex but not simple bursts, and second, by an interaction between the number of distractors and serial position for complex but not simple bursts. The shape of the predictions in Fig. 2 suggests that serial position be coded by its linear and quadratic contrasts: Any fanning of serial position curves as predicted by SOB for complex bursts would be reflected in a significant interaction between one or both of those contrasts of serial position and the number of distractors.[3] The effects corresponding to these core prediction of SOB (predictions 1a and 1b from above) are highlighted in the table.

As predicted by SOB, with simple bursts the number of distractors had no effect either on its own or in conjunction with any of the other variables. Again as predicted, there was a main effect of the number of distractors in the complex-burst group. Moreover, the predicted interaction between the number of distractors and serial position was significant for the quadratic (but not linear) contrast. To provide further support for the apparent differences in forgetting between simple and complex bursts, we conducted an omnibus ANOVA that included within-burst similarity as an additional between-subjects variable. That analyses yielded the crucial two-way interaction of number of dis-

[3] The fact that SOB predicts forgetting to increase across serial position in a negatively accelerated manner implies that the number of distractors should interact with the quadratic contrast of serial position. The further fact that this increase is not (or only slightly; see predictions for the complex-changing condition in Fig. 2) non-monotonic implies that the number of distractors may also interact with the *linear* contrast (for an illustration of the joint presence of linear and quadratic effects, see Keppel, 1991, p. 151). We therefore tested for the presence of both interaction contrasts in all experiments. We did not, however, consider the conventional unconstrained interaction between serial position and other variables because that interaction might, in the extreme case, arise even if the number of distractors had an effect only at a single serial position. Hence, the unconstrained interaction would be a very liberal test of the hypotheses under consideration, whereas coding serial position by its contrasts provides a more rigorous statistical test of SOB's predictions.

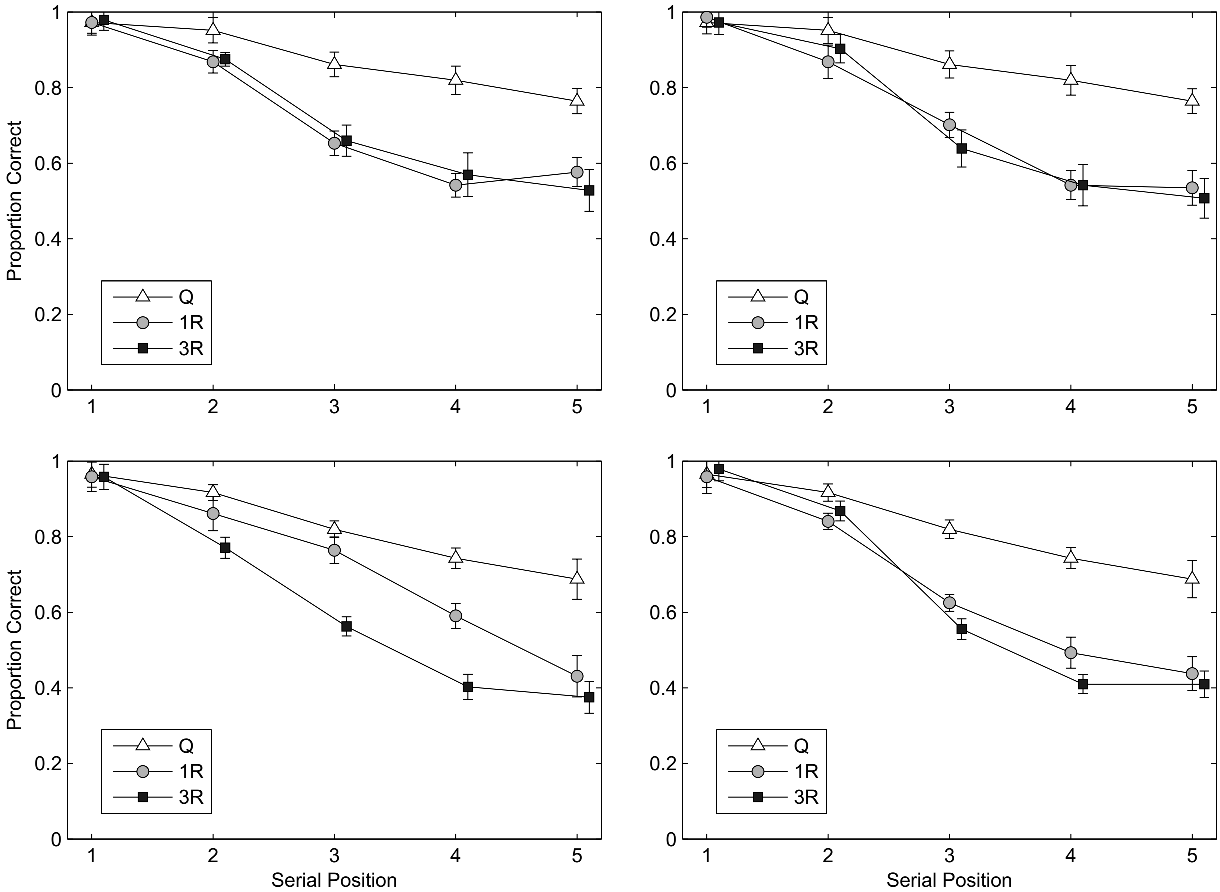

Fig. 3. Serial position curves obtained for all conditions in Experiment 1. The upper panels show data from the simple-burst group, the lower panels show data from the complex-burst group. Steady bursts are shown in the left-hand panels and changing bursts in the right-hand panels. The error bars represent standard errors.

tractors and within-burst similarity, $F(1,22) = 5.1$, $p < .05$, partial $\eta^2 = .19$, and the three-way interaction between number of distractors, within-burst similarity, and the quadratic contrast of serial position, $F(1,22) = 6.83$, $p < .02$, partial $\eta^2 = .24$, thus buttressing our conclusions that lengthening delays at retrieval causes additional forgetting only with complex bursts as predicted by SOB.

The second prediction of SOB, a small negative effect of changing compared to steady bursts, was not supported—the main effect of between-burst similarity was non-significant in both groups. The third prediction, that changing between-burst similarity should reduce the amount of forgetting compared to steady bursts, was borne out by the data: Between-burst similarity interacted with the number of distractors in the complex but not the simple group. In the omnibus analysis including both groups, the three-way interaction of between-burst similarity, number of distractors, and within-burst similarity was significant, $F(1,22) = 5.3$, $p = .03$, partial $\eta^2 = .19$.

Discussion

Experiment 1 in large part confirmed the predictions of SOB: there was no effect of the number of distractors and no fanning for simple bursts, irrespective of whether or not they changed across output positions (none of the effects involving number of distractors was significant for that group). By contrast, there was significant forgetting over filled delays with complex bursts, as indicated by the significant main effect of number of distractors and its interaction with serial position. The presence vs. absence of forgetting was moreover confirmed by the relevant interactions in the omnibus analysis involving all experimental variables.

The prediction regarding the relative size of the forgetting effects with steady vs. changing between-burst similarity was also at least partially confirmed. In the complex group, the effect of number of distractors was smaller with changing than with steady bursts, whereas between-burst similarity had no effect for simple distractors.

One feature of Experiment 1 that may have contributed to the outcome was the use of a well-memorized list of distractor items (months of the year). It is possible that this limited the effects of the distractors because recitation of three consecutive months from memory is likely to involve limited "novelty", given the extent to which those items have been rehearsed over a participant's life-time. We examined this possibility in the next study.

Table 2
Summary of ANOVAs on recall accuracy in Experiments 1 and 2

Source	Simple bursts			Complex bursts		
	F	p	Partial η^2	F	p	Partial η^2
Serial position (SP; L)						
Experiment 1	52.8	<.00	.83	90.2	<.00	.89
Experiment 2	66.8	<.001	.86	124.9	<.001	.92
Serial position (SP; Q)						
Experiment 1	6.54	<.05	.37	12.8	<.00	.54
Experiment 2	4.74	.05	.30	31.4	<.001	.74
No. of distractors (ND)						
Experiment 1	0.1	.75	.01	***17.6***	***<.001***	***.62***
Experiment 2	2.6	.14	.19	***22.2***	***.001***	***.67***
ND × SP (L)						
Experiment 1	0.9	.37	.07	***3.3***	***.09***	***.23***
Experiment 2	3.8	.08	.26	***5.4***	***.04***	***.33***
ND × SP (Q)						
Experiment 1	0.4	.56	.03	***8.1***	***<.02***	***.42***
Experiment 2	0.1	.77	.01	***3.1***	***.11***	***.22***
Between-burst similarity (BS)						
Experiment 1	0	.88	0	0.8	.39	.07
Experiment 2	0.1	.73	.01	1.3	.28	.10
BS × SP (L)						
Experiment 1	0.8	.39	.07	0.5	.49	.04
Experiment 2	0.1	.76	.01	2.1	.18	.16
BS × SP (Q)						
Experiment 1	0.6	.47	.05	2.6	.14	.19
Experiment 2	0.0	.95	.0	.77	.40	.07
BS × ND						
Experiment 1	0.2	.64	.02	7.4	.02	.40
Experiment 2	1.4	.27	.11	5.3	.04	.33
ND × BS × SP (L)						
Experiment 1	0	.90	0	0	1.0	0
Experiment 2	0.2	.65	.02	2.9	.12	.21
ND × BS × SP (Q)						
Experiment 1	0.5	.52	.04	2.7	.13	.20
Experiment 2	2.1	.18	.16	0	.98	0

Serial position is coded as linear (L) and quadratic (Q) contrast. All $df = 1, 11$. Effects of greatest theoretical interest (i.e., those relating to SOB predictions 1a and 1b) are bold-faced in italics.

Experiment 2

Experiment 2 extended the findings of Experiment 1 by using arbitrary distractor words whose sequence was not pre-experimentally memorized. Distractors consisted of a random list of common words taken from the MRC psycholinguistics database (Coltheart, 1981; Wilson, 1987). Within-burst similarity was again manipulated between subjects, whereas the remaining two variables (number of distractors, 1 vs. 3; and between-burst similarity, steady vs. changing) were manipulated within subjects.

Methods

Participants

Participants were 24 members of the University of Western Australia campus community. They participated voluntarily in a 1-h session for which they were remunerated with A$10. The participants were randomly assigned to one of the two experimental conditions.

Materials and procedure

The experiment was identical to Experiment 1 with the exception of the nature of the distractors. A pool of 200 distractors was sampled from the MRC psycholinguistics database. All words were nouns, consisted of 6–14 characters, 2–3 syllables, and had a written frequency >89 (Kucera & Francis, 1967). Pairs of participants were randomly yoked across conditions, and each pair received a different random selection of 12 words from the pool of 200.

As in Experiment 1, the same set of 12 words provided the distractors for a given subject across all trials. A complex burst consisted again of three different words (e.g., "office", "summer", "table") whereas the simple burst would repeat the same word (e.g., "office", "office", "office"). In the complex-changing condition, participants went through all 12 words in their set in a constant order on each trial, whereas in the other conditions only a random subset of those words was used (e.g., 4 in the simple-changing condition, and so on). The procedure was identical to that used in Experiment 1.

Results

Timing manipulation

Table 1 shows that the pattern of response times (again averaged across output positions 2–5) closely mirrored the recall times observed in Experiment 1. Compared to the baseline, the presence of a single distractor again slowed responding by about one second; $F(1,11) = 127.8$, $p < .0001$, partial $\eta^2 = .92$, in the simple-burst group, and $F(1,11) = 377.5$, $p < .0001$, partial $\eta^2 = .97$, in the complex-burst group. Increasing the number of distractors from 1 to 3 further slowed responding for both simple and complex bursts by about one second; $F(1,11) = 107.1$, $p < .001$, partial $\eta^2 = .91$, in the simple-burst group; and $F(1,11) = 78.7$, $p < .001$, partial $\eta^2 = .87$, in the complex-burst group.

Additional effects of between-burst similarity indicated that the changing nature of distractors further increased the articulation time and thus led to slower recall of subsequent items. Unlike in Experiment 1, changing bursts slowed recall not only in the complex-burst group, but also in the simple group; $F(1,11) = 20.7$, $p < .001$, partial $\eta^2 = .65$, and $F(1,11) = 19.5$, $p < .001$, partial $\eta^2 = .64$, for simple- and complex-bursts, respectively.

Recall performance. Correct-in-position recall performance ranged from .45 to .93 across individuals, and thus all participants were retained for analysis. As in Experiment 1, a 2 (within-burst similarity: simple vs. complex) × 5 (serial position) between-within ANOVA on performance in the baseline condition showed no main effect of within-burst similarity, nor an interaction with serial position (both Fs < 1).

Fig. 4 shows the serial position curves for all conditions in the usual layout of panels. The effects of our distractor-similarity manipulations were again first analyzed with two separate ANOVAs, whose results are shown in Table 2. In general, the effects from the first study were replicated here with the different distractor material.

Crucially, we replicated the main effect of the number of distractors in the complex-bursts group, although this time it was also accompanied by the interaction with the linear (rather than quadratic) trend across serial positions. Those effects were again absent in the simple-burst group. The combined pattern of the effects conformed to SOB's primary predictions about when forgetting should and should not be observed. In further support, the significant interaction of number of distractors with within-burst similarity in the omnibus between-within ANOVA, $F(1,22) = 9.4$, $p < .006$, partial $\eta^2 = .30$, confirmed that a larger number of distractors led to additional forgetting only when the bursts were complex. Unlike in Experiment 1, the three-way interaction involving within-burst similarity,

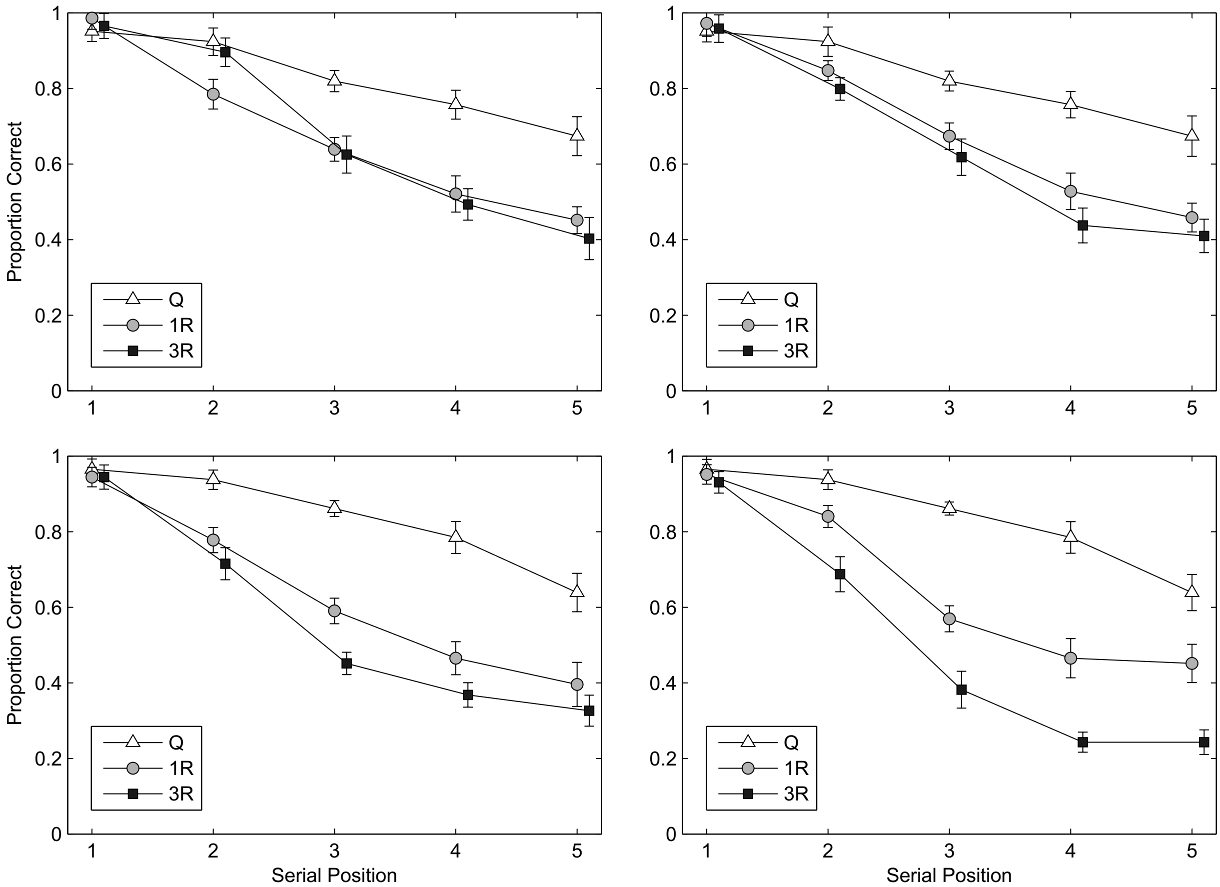

Fig. 4. Serial position curves obtained for all conditions in Experiment 2. The upper panels show data from the simple-burst group, the lower panels show data from the complex-burst group. Steady bursts are shown in the left-hand panels and changing bursts in the right-hand panels. The error bars represent standard errors.

number of distractors, and serial position failed to reach significance for both the linear, $F(1,22) = 1.02$, $p > .1$, partial $\eta^2 = .04$, and quadratic contrast, $F(1,22) = 2.66$, $p \cong .12$, partial $\eta^2 = .11$.

As in Experiment 1, we observed an interaction involving between-burst similarity and the number of distractors in the complex but not the simple group, although the serial position curves in Fig. 4 suggest that the interaction reflected a different pattern: It appears as though the effect of number of distractors in the complex group was *greater* with changing bursts than with steady bursts, whereas the reverse had occurred in Experiment 1. As in Experiment 1, neither the main effects of between-burst similarity nor the three-way interaction between serial position, number of distractors and between-burst similarity were significant in either analysis.

Discussion

Experiment 2 differed from Experiment 1 in the distractor materials used. Instead of a small and well-learned set (the months of the year), each participant received a different random sample of 12 items from a pool of 200 common English words. On the one hand, this change in distractor material did not affect the outcome because we again observed forgetting over filled delays with complex bursts but not with simple bursts (although the way in which this effect interacted with serial position was altered compared to Experiment 1). On the other hand, the change in material altered the effect of between-burst similarity in the complex group: whereas forgetting was reduced with changing bursts in Experiment 1, it increased compared to steady bursts in Experiment 2.

Experiment 3 addressed this inconsistency by providing a within-experiment comparison of the effects of distractor familiarity.

Experiment 3

Experiment 3 only involved complex bursts but again manipulated between-burst similarity and number of distractors within subjects. In addition, we examined the effects of the familiarity of distractor items between subjects: In the unfamiliar-distractor group, people received a random sample of words to articulate (as in Experiment 2), whereas in the familiar-distractor group people again articulated months of the year, as in Experiment 1.

Methods

Participants

Participants were 40 members of the University of Western Australia campus community. They participated voluntarily in a 1-h session for which they were remunerated with A$10.

Materials and procedure

Experiment 3 was a replication of the complex-burst procedure used in the first two experiments. Participants were randomly assigned to one of two groups ($N = 20$ in each group): One group articulated a highly familiar list of distractors as in Experiment 1 (months of the year), and the other group articulated a random sample of 12 words from a larger pool as in Experiment 2. We refer to this between-subjects variable as distractor familiarity.

In all other respects, the procedure was identical to that used for the complex-burst lists in Experiments 1 and 2.

Results

Timing manipulation

Table 1 reports the observed response times averaged across output positions 2–5. The insertion of a single distractor slowed recall by about 1 s in comparison to baseline for both the high-familiarity, $F(1,19) = 426.2$, $p < .001$, partial $\eta^2 = .96$, and the low-familiarity groups, $F(1,19) = 213.4$, $p < .0001$, partial $\eta^2 = .92$. Similarly, increasing the number of distractors from 1 to 3 slowed recall by a further 1 s for the high-familiarity group, $F(1,19) = 221.58$, $p < .0001$, partial $\eta^2 = .92$, and for the low-familiarity group, $F(1,19) = 210.23$, $p < .0001$, partial $\eta^2 = .92$. In addition, between-burst similarity had an effect for both groups, with changing bursts slowing recall compared to steady bursts for both the high-familiarity group, $F(1,19) = 27.32$, $p < .0001$, partial $\eta^2 = .59$, and the low-familiarity group, $F(1,19) = 19.52$, $p < .0003$, partial $\eta^2 = .51$. Although there was a trend for the low-familiarity group to take longer overall, the main effect of group was not significant, $F(1,38) = 2.7$, $p = .11$, partial $\eta^2 = .07$.

We conclude that varying the number of distractors between retrievals successfully manipulated retrieval time. Additionally, people took slightly more time to say different words in between recalls as compared to saying the same set of words over again.

Recall performance. Correct-in-position recall performance ranged from .41 to .92 across individuals and thus all participants were retained for analysis. We again compared recall performance of both groups in the baseline conditions. A 2 (distractor familiarity) $\times$ 5 (serial position) between-within ANOVA showed no main effect of group and no interaction (both Fs < 1).

Fig. 5 shows the serial position curves for all conditions. Note that unlike in the preceding experiments, all distractors in this experiment were complex. Rows of panels indicate familiarity with the distractors, with familiar distractors (months of the year) shown at the top and unfamiliar distractors (random words) at the bottom. As before, columns of panels refer to steady (left) and changing (right) between-burst similarity. The results of separate ANOVAs for each group are shown in Table 3. The table again shows the effects that reflect SOB's core predictions in bold.

As expected from the SOB predictions, and replicating the results with complex distractors in Experiments 1 and 2, the number of distractors had a significant main effect in both groups. Moreover, the number of distractors interacted with the trend of serial position in both groups. For familiar distractors, the linear-contrast interaction was significant, whereas for unfamiliar distractors it was the interaction with the quadratic trend that was significant. We conclude that irrespective of distractor familiarity,

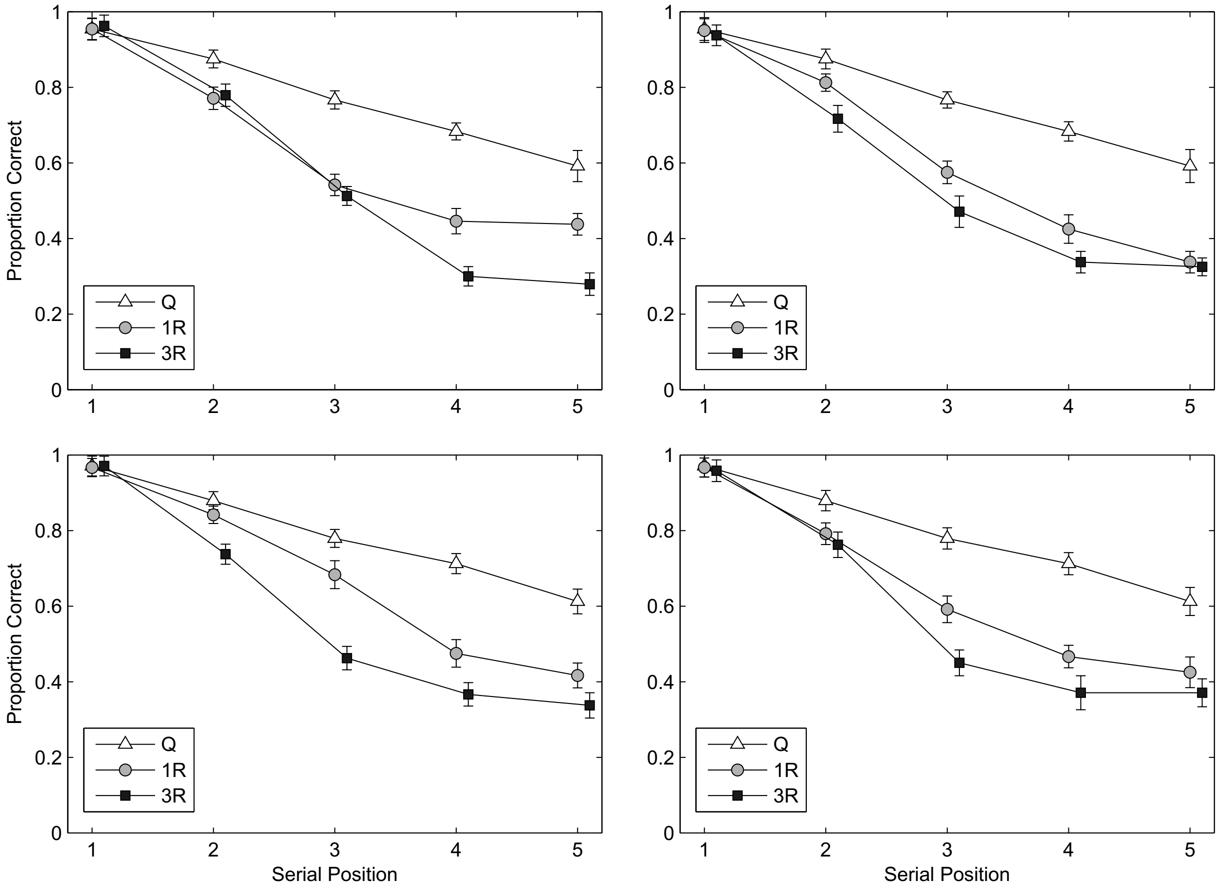

Fig. 5. Serial position curves obtained for all conditions in Experiment 3. The upper panels show data from the high-familiarity distractor group (paralleling complex bursts in Experiment 1), the lower panels show data from the low-familiarity distractor group (paralleling complex bursts in Experiment 2). Steady bursts are shown in the left-hand panels and changing bursts in the right-hand panels. The error bars represent standard errors.

Table 3
Summary of ANOVAs on recall accuracy in Experiment 3

Source	Familiar distractors			Unfamiliar distractors		
	F	p	Partial η^2	F	p	Partial η^2
Serial position (SP; L)	293.0	<.001	.94	152.0	<.001	.89
Serial position (SP; Q)	16.5	<.001	.47	36.4	<.001	.66
No. of distractors (ND)	***10.4***	***.004***	***.35***	***15.9***	***.001***	***.46***
ND × SP (L)	***11.4***	***.003***	***.38***	***3.3***	***.08***	***.15***
ND × SP (Q)	***2.1***	***.17***	***.10***	***13.2***	***.002***	***.41***
Between-burst similarity (BS)	0.4	.51	.02	0.5	.47	.03
BS × SP (L)	0	.94	.02	1.2	.28	.06
BS × SP (Q)	1.2	.28	.06	2.6	.12	.12
BS × ND	0	.98	0	2.3	.15	.11
ND × BS × SP (L)	9.4	.006	.33	0	.91	0
ND × BS × SP (Q)	8.57	.009	.31	1.9	.18	.09

Serial position is coded as linear (L) and quadratic (Q) contrast. All df = 1,19. Effects of greatest theoretical interest (i.e., those relating to SOB predictions 1a and 1b) are bold-faced in italics.

complex bursts caused fanning of the serial position curves; however, the balance between the linear and quadratic components of that fanning differed with distractor familiarity.

Our conclusion was buttressed by the omnibus between-within ANOVA, which yielded significant interactions of the number of distractors with the linear contrast of serial position, $F(1,38) = 12.4$, $p \approx .001$, partial $\eta^2 = .25$, as well with the quadratic contrast, $F(1,38) = 14.14$, $p < .001$, partial $\eta^2 = .27$. In addition, the highest-order interaction involving all four variables was also significant for the linear, $F(1,38) = 6.1$, $p < .02$, partial $\eta^2 = .14$, and quadratic, $F(1,38) = 9.11$, $p < .005$, partial $\eta^2 = .19$, contrasts, confirming that the extent of fanning was modulated by both between-burst similarity and familiarity with the distractors.

The interaction of between-burst similarity with the number of distractors, which went in opposite directions in the two preceding experiments, was non-significant in both groups in the present experiment. We conclude that these interactions are unreliable and we therefore do not discuss them further.

Discussion

Experiment 3 provided closure on the effects of distractor familiarity. Setting aside the volatile role of between-burst similarity, the pattern across Experiments 1–3 was quite consistent: highly familiar distractors are associated with less forgetting than unfamiliar distractors (i.e., comparison of the bottom panels of Figs. 3 and 4 suggests the same effect as the comparison of the top and bottom panels, respectively, of Fig. 5). Numerically, if the extent of forgetting is computed for all complex-burst conditions across the first three experiments (by taking the average difference across serial positions between 1 and 3 distractors), it turns out to be slightly smaller for familiar distractors (.06) than for unfamiliar distractors (.10). The slightly smaller disruptive effect of familiar distractors is compatible with the novelty-sensitive encoding postulated by SOB: The novelty associated with well-rehearsed sequences such as months of the year is likely to be lower than that of randomly-chosen words. In confirmation, we conducted an additional simulation (not reported in detail here) in which we modeled pre-experimental familiarity by amplifying the initial values of the connections in the weight matrix. This manipulation reduced the predicted extent of forgetting incurred by complex bursts by up to 3 percentage points—roughly comparable to the empirically observed reduction. We therefore do not consider the role of distractor familiarity further.

Turning to the principal result of Experiment 3, we again replicated the increased forgetting over longer filled delays that emerges with complex bursts, irrespective of whether or not the bursts change across output position and whether or not the distractors are highly familiar. This forgetting associated with complex bursts stands in marked contrast to its virtual absence in the first two experiments with simple distractors (and in the experiments reported by (Oberauer & Lewandowsky, in press). The outcomes of the first three experiments are therefore consistent with the predictions of SOB. However, before we can accept the interference mechanism postulated by SOB as an explanation of how forgetting can be experimentally turned on or off over identical time scales, we must consider a potential alternative that appeals to differences in cognitive load and the opportunity for attentional memory refreshing.

Specifically, there is evidence that repeated articulation of the same word requires full attention only for the brief time during which the speech plan is established (e.g., Naveh-Benjamin & Jonides, 1984); thereafter, some attentional capacity becomes available during articulation that may be devoted to other operations, including perhaps refreshing of memory traces. Several recent developments have suggested that verbal representations in short-term memory can be revitalized by means other than articulatory rehearsal; namely, by attentional refreshing (e.g., Barrouillet, Bernardin, & Camos, 2004; Oberauer & Lewandowsky, in press; Raye et al., 2007). Direct evidence for the notion of attentional refreshing was provided by Hudjetz and Oberauer (2007), who found that *slowing* of continuous reading of distractors during encoding *improved* memory performance—in other words, providing additional encoding time without adding additional interfering material while preventing articulatory rehearsal was associated with a memorial benefit. That benefit arguably arose from the additional attentional refreshing that was possible during slowed reading.

It follows that simple bursts may have permitted attentional refreshing of memory traces once the speech plan had been established, thus preserving memory not because novelty-sensitive interference was minimal but because the presence of decay was masked by compensatory refreshing. Complex bursts, by contrast, on the attentional-refreshing hypothesis would not have permitted refreshing because each new word required attention, thus providing an opportunity for decay to express itself.

This alternative account makes a readily testable prediction that permits its differentiation from the interference explanation offered by SOB: if simple bursts failed to cause forgetting in the first three experiments because they permitted concurrent attentional refreshing, then forgetting should emerge even with simple bursts if concurrent refreshing is blocked by another attentional task. Experiment 4 examined this possibility.

Experiment 4

The presumed attentional refreshing mechanism is conceptualized to involve a domain-general bottleneck that supports all central cognitive processes, including retrieval and response selection (Barrouillet et al., 2004). The existence of the bottleneck has been firmly established by research on dual-task performance, in particular with the psychological refractory period paradigm (Pashler, 1994; Pashler, Johnston, & Ruthruff, 2000; Ruthruff, Pashler, & Klaassen, 2001). Work with this paradigm has identified response selection in speeded choice tasks as requiring the bottleneck (Pashler, 1994)—the same bottleneck that has been linked with retrieval from long-term memory (Rickard & Bajic, 2004) and, important for the present context, encoding into short-term memory (Jolicoeur & Dell'Acqua, 1998).

It follows that if simple bursts are accompanied by a choice task, they place continuous demands on the attentional bottleneck, as is assumed for the complex bursts used in the first three experiments. Hence, if differences in attentional refreshing were responsible for our results thus far, simple bursts accompanied by a choice task should give rise to forgetting in the same way as complex bursts. By contrast, if the results of the first three experiments arose from novelty-sensitive interference, as predicted by SOB, then forgetting should be negligible with simple bursts notwithstanding the addition of a speeded choice task. SOB makes this prediction because the individual distractor events in each burst (i.e., combinations of saying the distractor word once, perceiving one of two possible visual stimuli, and executing the corresponding key press response) are all very similar to each other, thus producing little additional encoding for all but the first distractor event in each burst. To provide a particularly strong test of the hypothesis, the maximum number of distractor events in Experiment 4 was increased from 3 to 4, and the first burst preceded recall of the first item (rather than following it as in Experiments 1–3).

Methods

Participants

A sample of 20 psychology undergraduates at the University of Bristol participated voluntarily in exchange for course credit.

Materials and procedure

Participants were presented with lists of 5 consonants, assembled by drawing items at random without replacement from a pool of 19 consonants (excluding Q and Y). List presentation parameters were identical to the preceding experiments.

There were only two conditions, one with a single distractor event preceding each item at retrieval, and one with four distractor events. All bursts in this experiment were simple and involved articulation of the word "super". In contrast to the preceding studies, each distractor event additionally entailed a choice trial involving the symbols "&" and "%" that were arbitrarily mapped to response keys. For each distractor event, people had to press the left arrow key in response to the ampersand, and the right arrow key in response to the percentage symbol, in addition to articulating the distractor.

A cue indicating the condition was displayed immediately after offset of the last list item for 700 ms. Recall of each item was prompted by a question mark. Participants recalled each consonant orally and pressed the up-arrow key to proceed to the next burst of distractors which commenced 100 ms later. One or four distractor events preceded each recall event. When four distractor events were presented, they followed immediately upon each other (with a 100-ms blank interval between each key press and the next distractor stimulus).

Prior to commencing the main memory task, participants practiced the choice task on its own for 100 trials to ensure mastery of the arbitrary stimulus–response mapping and to avoid the need for participants to maintain a verbal description of the mapping in memory during the experiment proper. The memory task commenced with 20 practice trials, followed by 64 test trials. Trials were separated by a blank interval of 3.5 s. A short break was provided after every 20 trials. The experimental session lasted about 50 min.

Results

Correct-in-position recall performance ranged from .34 to .93 and thus all participants were retained for analysis. Fig. 6 shows that there was no convincing evidence for increased forgetting with four, compared to one, distractor events, as predicted by the decay hypothesis augmented by attentional refreshing. In a 2 (Delay; 1 vs. 4 distractor events) × 5 (Serial position) within-subjects ANOVA the

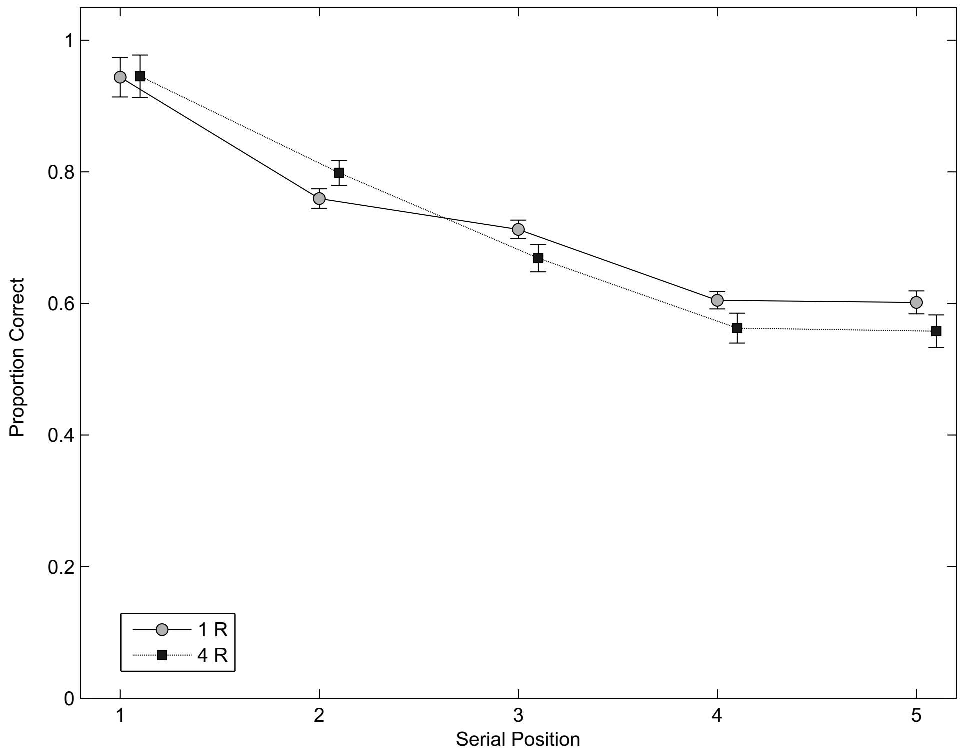

Fig. 6. Serial position curves for both delay conditions in Experiment 4. All bursts are simple and accompanied by a symbolic choice task. The error bars represent standard errors.

critical main effect of delay was not significant, $F(1, 19) = 1.8$, $p = .20$, partial $\eta^2 = .09$. Delay interacted with the linear trend of serial position, $F(1, 19) = 5.4$, $p = .031$, partial $\eta^2 = .22$, but not with the quadratic contrast, $F(1, 19) = .01$, $p > .10$, partial $\eta^2 = .09$. However, the linear interaction did not reflect fanning from a common origin. Rather, the interaction was driven by a crossing over of the serial position curves such that the condition with longer delay performed *better* at early serial positions and only slightly worse at later serial positions: there was no indication that longer delays were consistently associated with more forgetting at any serial position. In confirmation, separate *t*-tests comparing 1 and 4 distractors for the five serial positions revealed that none of them differed significantly—from first to last, $t(19) = -.19$, -1.98, 1.89, 1.88, 1.47, respectively (the corresponding *p* values were .85, .06, .07, .08, and .16).

Accuracies for the choice task were .95 in the condition with one distractor event and .96 in the condition with four distractor events. After removal of 11 outliers exceeding 20 s, mean latencies at output positions 2–5 were computed as 2.46 s for recall of an item and one preceding distractor event, and 5.09 s for recall of an item together with four preceding distractor events. Table 1 shows that those times spanned a considerably wider range than that covered by the first three experiments.

Discussion

Experiment 4 provided no support for the hypothesis that attentional refreshing could have prevented forgetting in the simple-burst conditions: notwithstanding the considerably larger time scale and notwithstanding the addition of a demanding attentional task as well as an additional burst of distractors before retrieval commenced, the simple bursts were not associated with any detectable forgetting. This result replicates and extends the related finding reported by Oberauer and Lewandowsky (in press).

It follows that the lack of forgetting with simple bursts in Experiments 1, 2, and 4 did not reflect the inadvertent masking of temporal decay by attentional refreshing. By implication, the differences between simple and complex bursts that were observed across all four studies also cannot be explained by the attentional refreshing hypothesis.

It could conceivably be argued that the choice task in Experiment 4 was less successful in blocking attentional refreshing than the complex bursts in the first three experiments. There are several difficulties with this argument: first, there is no evidence or theory to suggest that reading of three different words is attentionally *more* demanding than four consecutive symbolic choice tasks with arbitrary stimulus–response mappings. Second, it is not clear why two tasks performed concurrently—articulation and choice—should block attentional processing *less* than reading of three words on its own. Third, the combined distractor events used in Experiment 4 took considerably more time than reading a word in the complex bursts of Experiments 1–3; this observation further militates against the assumption that the combined distractor effects of Experiment 4 were less attention demanding. We therefore do not consider the attentional refreshing hypothesis any further and instead conclude that the differences between simple and complex bursts are best explained within the novelty-sensitive interference framework embodied in SOB.

The cross-over interaction of number of distractors with serial position was unexpected but should not easily be dismissed as spurious. A similar cross-over pattern was observed on several occasions in the previous experiments, most notably in the simple-burst groups of Experiments 1 and 2. This pattern suggests that even when an increased delay has no effect on overall accuracy (in no case was there a significant main effect of the number of distractors or a significant disadvantage of additional delay for later list items), it may occasionally tilt the serial position curve towards more primacy. No existing model of serial recall predicts that recall gets better in the primacy portion when it is delayed by a distractor-filled interval. The effect is reminiscent of the recency-to-primacy shift that is sometimes obtained with increased retention intervals in short-term recognition tasks (Knoedler, Hellwig, & Neath, 1999; Neath, 1993). A joint investigation of serial recall and recognition might be a fruitful avenue to explore this effect further.

General discussion

Summary of results

Across the four experiments, we observed two distinct patterns of forgetting that conformed precisely to the "now you see it, now you don't" pattern that has beset the previous literature. We sometimes observed forgetting as a function of increased delay and we sometimes did not.

In contrast to the previous literature, however, the two distinct patterns obtained here were under complete experimental control and in line with the predictions of SOB presented at the outset: When people repeatedly articulated the same word within each burst, no additional forgetting was observed if the duration of the distractor task was extended. By contrast, when people articulated different words within each burst, extending the distractor activity caused additional forgetting. Within-burst similarity is crucial and between-burst similarity has no, or at best a small, effect on the extent of forgetting. Statistical support for this over-arching conclusion was provided by an omnibus between-within ANOVA that considered all participants from all four studies together (to render the designs commensurate, this analysis collapsed across between-burst similarity for Experiments 1–3). This analysis revealed the diagnostic two-way interaction between the number of distractors and within-burst similarity, $F(1, 106) = 17.95$, $p < .0001$, partial $\eta^2 = .14$, as well as the predicted three-way interaction between the number of distractors, within-burst similarity, and the quadratic contrast of serial position, $F(1, 106) = 14.08$, $p < .0003$, partial $\eta^2 = .12$ (Another analysis, restricted to the first three experiments, included between-burst similarity as a factor and found that none of the effects associated with that variable reached significance.)

To provide an over-arching perspective on our data, Fig. 7 presents the accuracy serial position curves (top panels) together with the underlying cumulative retrieval

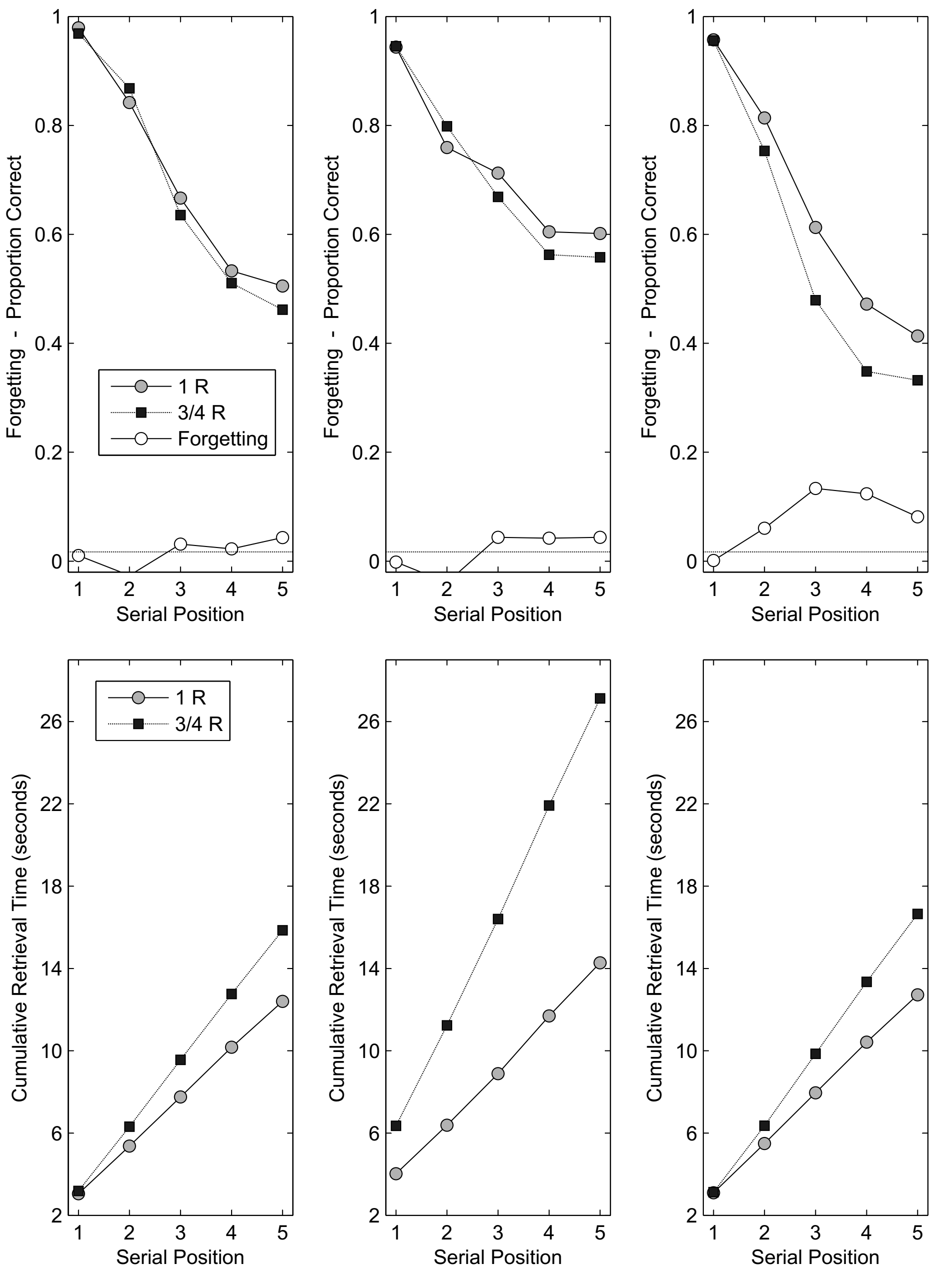

Fig. 7. Summary of results from all four experiments, averaging across the familiarity of the distractors and between-burst similarity. The top panels show accuracy serial position curves and the bottom panels show the accompanying cumulative retrieval times. The top panels also show the amount of forgetting by plotting the differences between short and long delays (the horizontal lines show the average forgetting with simple bursts across Experiments 1, 2, and 4 as a reference point). The left column of panels shows the results with simple bursts of Experiments 1 and 2, averaging across experiment. The center panels show the data of Experiment 4, involving simple bursts accompanied by a speeded choice task. The panels on the right show results involving complex bursts in Experiments 1, 2, and both groups of Experiment 3, again averaging across experiment.

times (bottom panels) across all experiments. The data involving simple bursts in Experiments 1 and 2 are shown in the left-most column of panels, the data from Experiment 4 are shown in the center, and the data from all complex bursts in Experiments 1–3 on the right. The top panels also show the differences between the two serial position curves to illustrate the extent of forgetting due to the delay manipulation. To provide a reference point, the horizontal line shows the average forgetting across serial positions that was observed with simple bursts in Experiments 1, 2, and 4 combined (.017, i.e., less than 2 percentage points).

The figure clarifies that the increase in forgetting that resulted when bursts changed from simple to complex was not driven by a change in cumulative retrieval time (cumulative retrieval time is the time from offset of the last list item to the key press indicating completion of recall in a given serial position). Quite on the contrary, the temporal consequences of the distractor manipulation were largest in Experiment 4 (center panels), and yet that study exhibited considerably less forgetting than the complex bursts in the other three experiments with a lesser temporal manipulation (right-hand panels).

By implication, whatever caused the differential effects of distractors on forgetting, it was not the passage of time per se: no notable additional forgetting occurred with simple bursts in Experiment 4 when recall of the last item was delayed by an additional 13 s, whereas performance was impaired by some 10% with complex bursts when recall of the last item was delayed by only 4 additional seconds. Across serial positions, the forgetting observed with complex bursts was more than four times greater (.08) than that observed with simple bursts.

A final summary of our data is given by Fig. 8, which is a state-trace plot of the results from all four experiments. State traces are a tool for demonstrating dissociations between two variables or processes (Bamber, 1979; Loftus, Oberg, & Dillon, 2004). Fig. 8 shows, for each experiment and each serial position, the mean error rate with long delays as a function of the corresponding error rate with short delays, separately for simple bursts (open circles and squares) and complex bursts (filled circles), but averaging across other variables (e.g., between-burst similarity) where present. The logic of state-trace analysis is as follows: if forgetting were driven by a single common process, then a single monotonically increasing function must describe all data points for simple and complex bursts. By contrast, if forgetting differed with the type of distractor burst, then the data points from the two types of bursts would be better described by separate functions than by a joint function. Inspection of the figure strongly suggests that the two classes of data points fall onto two separate functions, with complex bursts leading to more forgetting (steeper slope) than simple bursts. Disregarding the non-diagnostic points crowded around the origin, there is only a single instance in which a complex burst led to less forgetting than a comparable simple burst (i.e., the filled circle on the dashed line towards the top of the graph).

Moreover, the figure once again clarifies that time per se does not determine forgetting: time (represented by the response times for the conditions with long delays) is proportional to the size of the plotting symbols: It is evident that the points for the simple bursts all stay close to the main diagonal, notwithstanding their considerable difference in size (i.e., response time) and notwithstanding

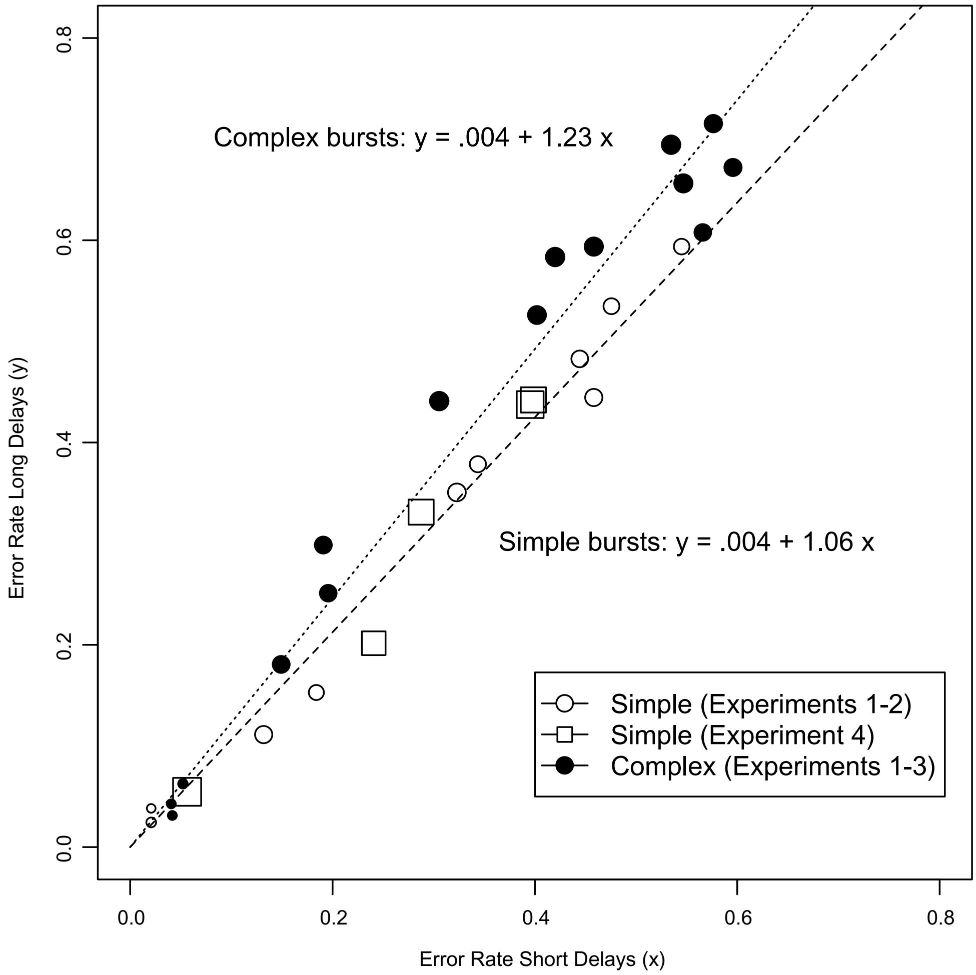

Fig. 8. State traces of simple and complex bursts. Error rates after long delays are plotted as a function of error rates after short delays for all serial positions in all experiments, aggregating across other experimental variables when present. Data points come from Experiments 1, 2, and 4 (simple bursts), and from Experiments 1, 2, and 3 (complex bursts). The two lines represent the two best-fitting slope estimates for simple and complex bursts, respectively. Size of plotting symbols is proportional to response time in the conditions with long delays (i.e., either 3 or 4 distractors, depending on experiment). Experiment 4 is represented separately to illustrate that its greater range of response times does not increase forgetting. See text for details.

the fact that some of those distractors were accompanied by an attention-demanding choice task (i.e., in Experiment 4, identified by the open squares).

Given that the plot showed no appreciable deviation from linearity, we fit two competing linear models to the data points. The single-function model simply uses short-delay error rates to predict long-delay error rates; this model yielded $R^2 = .957$. The dual-function model included an additional interaction predictor, viz. the product of short-delay error rates and within-burst similarity (simple = 0, complex = 1; this is isomorphic to postulating two separate slopes for the two types of burst). This model had a better fit, $R^2 = .975$, and the interaction predictor was significant, $F(1,27) = 18.9$, $p < 001$.

We compared these R^2 values using the likelihood ratio approach suggested by Glover and Dixon (2004, Eq. 2). If all observations in Fig. 8 are considered to be independent (hence $N = 30$), the likelihood ratio was 2923 in favor of the dual-function model, implying that the data were nearly 3000 times more likely under the dual-function model than under the single-function model.[4] This likelihood ratio provides further definitive evidence for a dissociation between simple and complex bursts with regard to their effect on error rates: With complex bursts, lengthening of delays causes more forgetting whereas with simple bursts it does not. We consider the empirical and theoretical implications of our results after we deal with potential criticisms that might be leveled at our studies.

Potential limitations and criticisms

One clear limitation of the present studies is that we restricted consideration to the verbal domain. All list items and all distractors consisted of verbal material, and our conclusions therefore do not necessarily extend to other domains, such as spatial tasks or stimuli. In particular, it is entirely unclear whether the within-burst similarity effects observed here would also arise with spatial interference tasks, such as tapping. Although spatial tasks can interfere with verbal memoranda (e.g., Guérard & Tremblay, in press), it remains to be ascertained by future research whether that interference is modulated by the similarity among spatial distractors.

Another potential criticism of the present studies might point out that our conclusions appear to rest, in part, on the absence of forgetting with simple bursts. By implication, it might be argued that we rely on acceptance of a null result. We counter that criticism in two ways. First, the near-absence of forgetting with simple bursts has by now been shown in about 10 experiments involving hundreds of participants (Lewandowsky et al., 2004; Oberauer & Lewandowsky, in press). In a meta-analysis of the available data, Oberauer and Lewandowsky found that overall performance loss was less than one-half of a percentage point (.0048) per second additional delay (95% confidence interval .0075 to .0021). The size of this effect, though statistically different from zero, is incommensurate with any available conceptualization of temporal decay. It follows that the negligible amount of forgetting observed with simple bursts in the present studies is unlikely to reflect a lack of statistical power that prevented the discovery of larger extents of forgetting—on the basis of the meta-analysis, increased statistical power would most likely yield a narrow confidence interval for the magnitude of forgetting, which may just fail to bracket zero, without however revealing sufficient forgetting to be compatible with decay as a major cause of the fragility of short-term memory traces.

Our second response underscores that the principal conclusions in this article rest on there being *more* forgetting with complex than with simple bursts. Under certain parameter values SOB expects some (limited) forgetting even with simple bursts (Oberauer & Lewandowsky, in press; see their Figs. 1 and 3). Moreover, the assumption—implemented in our simulations of SOB—that individual distractors in simple bursts give rise to exactly identical representations is unlikely to be perfectly accurate. Small variations in intonation of the same distractor word, for instance, would introduce a small amount of novelty even in simple bursts, leading so a small amount of additional interference. Therefore, whether or not simple bursts give rise to small and statistically significant forgetting is not terribly relevant: SOB necessarily predicts more forgetting with complex bursts, and across all experiments, our analyses have indubitably confirmed this prediction. Those analyses should put to rest any concerns about acceptance of a null result.

A final concern might focus on the quantitative agreement between our results and the predictions of SOB. Although the overall extent of predicted forgetting for simple and complex bursts (.003 and .081, respectively) was closely paralleled by the data (.017 and .080), there arguably was some deviation between predictions and data at the level of individual serial positions. Specifically, whereas there was a notable downturn in the observed extent of forgetting with complex bursts towards the end of the list (see open circles in the top-right panel of Fig. 7), this trend was smaller or absent in the predictions (see Fig. 2). Our response is twofold: first, the predictions of SOB were based on a priori parameter values derived from a different setting without consideration of the present data. It is therefore unreasonable to expect a close quantitative match between all aspects of predictions and data. Second, and perhaps most important, we know of no other existing theory that could even qualitatively—let alone quantitatively—predict the results observed in this article without considerable modification. In particular, purely temporal models (Burgess & Hitch, 1999; Brown, Neath, & Chater, 2007; Page & Norris, 1998) have no mechanism to accommodate the present time-independent effects of

[4] Because the regression involved repeated observations from the same participants (i.e., across serial positions within an experiment), the independence assumption may not be warranted. We therefore repeated the analysis using a Generalized Estimating Equations (GEE) approach (Hanley, Negassa, Edwardes, & Forrester, 2003). The GEE approach takes into account the correlation among observations within an experimental unit (in this case among serial positions in an experiment). Using the *geepack* within R (Højsgaard, Halekoh, & Yan, 2006), the GEE parameter estimates for the two-slope model were found to differ from the conventional regression estimates only in the third decimal place, suggesting that violation of the independence assumptions did not play a major role in this analysis.

within-burst similarity. Likewise, time-based models have no mechanism to accommodate the downturn in complex-distractor induced forgetting later in the list (see Fig. 1 in Lewandowsky et al., 2004), unlike SOB which at least sometimes predicts a (small) downturn; see Fig. 2. In consequence, a quantitative analysis of SOB's performance must await the development of potential contenders before a rigorous model comparison (e.g., Lewandowsky & Farrell, 2008; Oberauer & Lewandowsky, in press) becomes possible.

Empirical implications: reconciliations and links

The present article served to reconcile some seemingly contradictory results concerning forgetting in short-term memory under a common theoretical umbrella and within a common methodology. For example, the present studies reconciled the results of Page (2006) with those of Lewandowsky et al. (2004) and Oberauer and Lewandowsky (in press). The former used complex bursts of distractors and found substantial forgetting, whereas the latter two used simple bursts and found no or very little forgetting.

We next place our results into a larger context by revisiting the Brown–Peterson paradigm (e.g., Peterson & Peterson, 1959), which has been the main workhorse over several decades for studying sources of forgetting. In contrast to the present studies, the Brown–Peterson paradigm involves a single extended period of distractor activity that follows presentation of the memoranda, and once this period is over, recall proceeds without any further interruptions. A seemingly pervasive attribute of the Brown–Peterson paradigm is the massive and rapid forgetting that is observed as the distractor period is extended (e.g., Chechile, 1987; Peterson & Peterson, 1959). It may therefore be tempting to suggest that the present results do not transfer to the Brown–Peterson paradigm and are thus of limited generality. This impression, however, is incorrect. In the Brown–Peterson paradigm, just like in the present studies, forgetting is minimal or absent if the retention interval is filled with a simple burst of distractors. For example, Vallar and Baddeley (1982) obtained very little forgetting during a retention interval of up to 15 s that was filled by repeated articulation of a single word. Likewise, Phaf and Wolters (1993), using a similar paradigm and a retention interval filled with repeating the same distractor word over and over (i.e., a long simple burst) found small and generally non-significant forgetting over up to 60 s. In summary, contrary to common textbook summaries, the Brown–Peterson paradigm does not necessarily produce a declining forgetting curve over time, and whether or not it does depends on the novelty of each distractor, relative to the immediately preceding material.

We are not claiming that the dissimilarity among distractors is the only factor that determines how much forgetting occurs in a filled retention interval. One other factor that is likely to affect memory is the difficulty of the distractor activity (Barrouillet et al., 2004; for an early demonstration see Posner & Rossman, 1965). Indeed, in the Brown–Peterson paradigm, the distractors that give rise to forgetting (e.g., counting backwards by threes from a random number) are not only complex by our definition, but also arguably more difficult than repeated overt articulation of the same word. One question that arises from these precedents is whether our within-burst similarity effects might alternatively be explained by difficulty: Perhaps complex bursts are somehow more difficult than simple bursts.

We consider this unlikely for two reasons. First, in Experiment 1 distractors involved a well-learned canonical sequence, namely months of the year. Research into maintenance rehearsal has shown that repetitive speaking of even novel sequences is automatized very rapidly (cf. Aldrich, Garcia, & Mena, 1987), and reciting a well-learned sequence should be even less demanding. It follows that the complex bursts in Experiment 1 were unlikely to differ much in difficulty from simple bursts. Second, the simple bursts in Experiment 4 were accompanied by a second distractor task, involving choice based on arbitrary stimulus–response mappings. In addition to preventing attentional refreshing, this task also necessarily increased the overall difficulty of the distractor activity. Nonetheless, no effect of delay on forgetting was observed. Thus, the only way in which the difficulty hypothesis could accommodate the present data is by postulating that saying "April–May–June" is more difficult than saying "super" four times in a row while concurrently making four response choices involving arbitrarily mapped keys in response to "&" or "%", despite the fact that the latter distractor activity took much longer. We do not consider this credible.

Theoretical implications

Interference-based forgetting in SOB

The effects of within-burst similarity observed in our experiments were predicted by SOB at the outset without any adjustment to its free parameters, solely based on previous applications of its novelty-sensitive encoding mechanism. We consider this to constitute a strong *a priori* prediction for three reasons: first, at an architectural level, we used the definition of novelty that has been part of SOB since its inception (Farrell & Lewandowsky, 2002). Second, at the instantiation level, we exactly followed published precedent when deriving the predictions, by associating distractors to context markers (Oberauer & Lewandowsky, in press). Finally, novelty was instantiated and the predictions were obtained without knowledge of, or reference to, the present data.

Our results extend the conclusions of Oberauer and Lewandowsky (in press), who found in a quantitative comparison of various computational models that SOB provided a better quantitative account of forgetting data from two experiments than a decay-based model (the Primacy Model; Page & Norris, 1998) and a temporal-distinctiveness theory (SIMPLE; Brown et al., 2007). Those experiments all involved simple bursts of distractors and hence yielded virtually no forgetting when the number of distractors was increased from one to three or four. One principal theoretical contribution of this article is to show that SOB can simultaneously predict the much greater forgetting observed with complex distractors without any change in parameter values.

SOB correctly predicts that within-burst similarity is the main determinant of forgetting over a filled delay, but the predictions in Fig. 1 also show a small effect of between-burst similarity, which was not evident in the data. The effect of between-burst similarity arises in SOB from the overlap of context markers, such that distractors at later positions are regarded as less novel when they match distractors encoded at earlier positions (i.e., bursts are steady). In the context of SOB or any other model invoking novelty-sensitive encoding of distractors, our failure to find evidence for a role of between-burst similarity implies that novelty is computed in a highly ''local'' way, with little generalization across serial positions. The data suggest that a distractor repeated at a different serial position is regarded as novel and therefore encoded with full strength. SOB can accommodate a strongly local computation of novelty by assuming little overlap between successive context markers.

Forgetting or no forgetting: decay vs. rehearsal?

Whenever forgetting is found to be absent in a short-term memory task, decay advocates may argue that temporal decay was nonetheless present but masked by compensatory rehearsal. For example, Cowan and AuBuchon (2008) claimed that although Lewandowsky et al. (2004) successfully suppressed articulatory rehearsal with their simple bursts, they may have left attentional refreshing intact (cf. Hudjetz & Oberauer, 2007). The presence of attentional refreshing, in turn, may have masked temporal decay in the experiments of Lewandowsky et al. (2004), and the simple-burst conditions of the present Experiments 1 and 2. However, the present Experiment 4 provided strong evidence against this possibility, as did additional experiments by Oberauer and Lewandowsky (in press). Therefore, we no longer consider the hypothesis that decay has been masked by attentional rehearsal a strong contender.

Instead, we conclude this article by considering whether *any* type of memory refreshing, be it articulatory rehearsal, attentional refreshing, or a yet to-be-specified third mode that can survive the dual-distractor methodology of Experiment 4, permits reconciliation of the decay notion with the absence of forgetting. Perhaps somewhat surprisingly, the answer appears to be no. Oberauer and Lewandowsky (in press) explored the decay and rehearsal notions within the Primacy Model (Page & Norris, 1998) and within a model involving positional markers (inspired in large part by the model of Burgess & Hitch, 1999). Perhaps contrary to intuition, the simulations showed that rehearsal during retrieval was not necessarily beneficial and under certain conditions even *harmed* performance. The observed harmful effects arose because rehearsal necessarily involves (covert) recall, and when this recall is erroneous, an erroneous list is re-encoded and strengthened. Retrieval errors during rehearsal thus cumulate, with inevitable adverse effects on accuracy that typically outweigh the potential benefits associated with memory restoration (One obvious exception involves situations in which performance is perfect, in which case rehearsal is not harmful but also not needed). It follows that the appeal to compensatory rehearsal or refreshing does not, in fact, enable decay models to accommodate situations in which distractor-filled intervals fail to produce forgetting.

Acknowledgments

Preparation of this paper was facilitated by a Discovery Grant and a Linkage International Grant from the Australian Research Council to the first and the third author, and an Australian Professorial Fellowship to the first author. Supplemental material at web page: http://www.cogsciwa.com.

Appendix A. Formal summary of SOB and simulations for predictions

SOB rests on three principal architectural commitments: (1) novelty-sensitive encoding as described in the text; (2) implementation of response suppression by the same process that governs encoding; and (3) dynamic iterative deblurring of partial images retrieved from memory. Following other recent applications of SOB (e.g., Farrell, 2006), we retained the first two commitments but relaxed the third one owing to technical limitations (i.e., excessive computing demands associated with dynamic deblurring). The complete source code of SOB (in MatLab) used to generate the present predictions is available at the web page provided in the Acknowledgments.

Architecture

The network consists of two layers; an input layer ($N = 16$) used to represent positional markers, and an output layer ($N = 150$) representing list items, where N refers to the number of units in each layer. The two layers are fully interconnected by a weight matrix $\mathbf{W}$.

Positional markers

SOB incorporates a positional marker for each list position, as is frequently assumed by models of memory (e.g., Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999). The similarity between any two positional markers is an exponential function of their absolute separation in list positions; that is:

$$\cos\left(\mathbf{p}_i, \mathbf{p}_j\right) = t_c^{(|i-j|)}, \quad \text{(A1)}$$

where i and j are the positions of the ith and j th items, $\mathbf{p}_i$ and $\mathbf{p}_j$ are distributed vectors representing positional markers at those positions, and t_c is a fixed parameter determining the overlap between successive markers ($t_c = 0.5$). By implication, SOB represents only positional information and is unaware of the temporal separation between items at encoding. This is in line with recent findings that has shown temporal isolation to be irrelevant in forward serial recall (e.g., Lewandowsky et al., 2006; Nimmo & Lewandowsky, 2005).

Encoding

During list presentation, items are associated with successive positional markers via standard Hebbian learning (see, e.g., Anderson, 1995):

$$\Delta \mathbf{W}_i = \eta_e(i)\mathbf{v}_i\mathbf{p}_i^{\mathrm{T}}, \quad \text{(A2)}$$

where $\mathbf{W}$ is the weight matrix connecting markers to items, $\mathbf{v}$ is the vector representing the ith presented item, and $\mathbf{p}$ is the positional marker for the ith serial position. The learning rate η_e for the ith association, $\eta_e(i)$, was calculated using the *energy* between the to-be-learned association and the information captured by $\mathbf{W}$ up to that point:

$$\eta_e(i) = \begin{cases} 1, & i = 1 \\ -\phi_e/E_i, & i > 1 \end{cases}, \quad \text{(A3)}$$

where ϕ_e is a free parameter, and E_i, the energy of the ith association, is given by:

$$E_i = -\mathbf{v}_i^{\mathrm{T}}\mathbf{W}_{i-1}\mathbf{p}_i. \quad \text{(A4)}$$

(see Farrell & Lewandowsky, 2002). The use of energy to weight incoming information instantiates the core "novelty-sensitive encoding" principle of SOB that is at the heart of its predictions for the present experiments. Energy is a measure of the consistency between novel information and what has already been learned or, in psychological terms, of the novelty of an item in the context of a particular positional marker. By reducing the encoding strength of items that are not entirely novel, energy-gating limits redundancy and, by implication, enhances memory for unique information (e.g., Lewandowsky & Farrell, 2008). In the present context, novelty-sensitive encoding also modulates the adverse effects of additional distractors.

Item representation

The use of novelty-sensitive encoding mandates an accurate representation of the similarity structure among input items. Accordingly, all study items in the present simulation were derived from a common item prototype (a binary vector that was randomly sampled for each replication) by retaining each feature with probability p_c and resampling it (from a symmetric binary distribution) with probability $1 - p_c$. The value of p_c was fixed to .7; this value was chosen to approximate the average similarity among vectors in a recent application of SOB (Lewandowsky & Farrell, 2008) that instantiated a behavioral multi-dimensional scaling solution of the actual stimuli.

Retrieval

Retrieval in SOB consists of stepping through the positional markers to consecutively cue for their associated items. For position i, retrieval is cued by placing positional marker $\mathbf{p}_i$ across the input layer, and updating the item unit activations:

$$\mathbf{v}_i' = \mathbf{W}_i\mathbf{p}_i, \quad \text{(A5)}$$

where the resultant vector, $\mathbf{v}_i'$, is a "noisy" version of $\mathbf{v}_i$, containing a blend of $\mathbf{v}_i$ and the other items on the list owing to the overlap among positional markers. The probability of recalling any item $\mathbf{v}_v$ is obtained by matching $\mathbf{v}_i'$ to the experimental vocabulary:

$$P(\mathbf{v}_v) = \frac{s(\mathbf{v}_i\prime, \mathbf{v}_v)}{\sum_{k=1}^{n} s(\mathbf{v}_i', \mathbf{v}_k)}, \quad \text{(A6)}$$

where n is the number of items in the experimental vocabulary and the function s, the similarity between the noisy vector and an item vector, is calculated as:

$$s(\mathbf{v}_i', \mathbf{v}_k) = \exp\left[-cD(\mathbf{v}_i', \mathbf{v}_k)^2\right]. \quad \text{(A7)}$$

The Euclidian distance measure D is weighted by the free parameter c, and—for computational reasons—is normalized by subtracting the minimum distance, across all the n items, from the distance for each item. A discrete response is then determined by random sampling according to the preceding equation.

The recalled item is then suppressed by adjusting the weights between the positional marker layer and the item layer, according to:

$$\Delta \mathbf{W}_j = \eta_s(j)\mathbf{v}_{o,j}\mathbf{p}_j^{\mathrm{T}}, \quad \text{(A8)}$$

where j is the output position, and $\mathbf{v}_{o,j}$ is the item recalled at the jth position.

To ensure that the extent of response suppression approximately matches that of learning, the learning rate for suppression, $\eta_s(j)$, is also determined from the energy of the recalled item j according to:

$$E_j = -\mathbf{v}_{o,j}^{\mathrm{T}}\mathbf{W}_{j-1}\mathbf{p}_j. \quad \text{(A9)}$$

The response suppression rate is given by:

$$\eta_s(j) = \frac{-E_j}{\phi_s E_1}, \quad \text{(A10)}$$

where ϕ_s is a free parameter, and E_1 is the energy of the item recalled at the first output position (see Farrell & Lewandowsky, 2002). Suppression thus uses the same Hebbian learning used at encoding, except that the negative learning rate (note the minus sign in the numerator of the preceding equation) implements "anti-learning" (Anderson, 1991) and hence attenuates an item's representation.

Reflecting a fairly general consensus in the short-term memory literature (e.g., Brown et al., 2000; Cowan, Saults, Elliott, & Moreno, 2002; Fitzgerald & Broadbent, 1985; Oberauer, 2003), retrieval of an item entails output interference. Output interference is modeled by adding Gaussian noise with standard deviation N_O (a free parameter) to each weight in $\mathbf{W}$ after each retrieval.

Distractor representation

Distractors, when present, were associated with the positional marker used for the immediately preceding retrieval and were encoded in the same way as items during study. All distractors within a burst were associated to the same marker.

Distractors were created as follows. In a first step, three distractor prototypes were independently generated from the item prototype by retaining each feature of the item prototype with probability s_c and resampling it (from a symmetric binary distribution) with probability $1 - s_c$. To reflect the obvious fact that distractors differed more from the ensemble of list items than they differed among each other, the value of s_c was fixed to half of the value of p_c (i.e., $s_c = .35$). Actual distractors were then derived from those distractor prototypes at each output position by retaining features of the prototype with probability J_c and resampling them with probability $1 - J_c$. For simple bursts, all distractors within a burst were identical copies of a single derivative from one of the prototypes. For complex bursts, by contrast, each distractor within a burst was independently derived from one of the three prototypes (For bursts of a single distractor, simple and complex did not differ and only one prototype was used in each case.).

Between-burst similarity was implemented as follows: For changing bursts, J_c was set to 0, which effectively implied no similarity of distractors between bursts. For steady bursts, J_c was set to .31, which represents its best-fitting estimate obtained by Oberauer and Lewandowsky (in press). The use of similar but not identical distractors even when bursts were steady reflected the assumption that although people nominally uttered the same word(s) across output positions, the internal representation of those words likely changed somewhat across serial positions (The predictions are not materially altered if J_c is set to unity for steady bursts.).

The predictions of SOB were thus based on five free parameters (ϕ_e, ϕ_s, N_O, c, and J_c for steady bursts) whose values were set to the best-fitting estimates obtained by Oberauer and Lewandowsky (in press, their Experiment 2). In addition, SOB used the fixed parameters $t_c = 0.5$, $p_c = 0.7$, and $s_c = 0.35$. The predictions shown in Fig. 1 were obtained with $\phi_e = 300$, $\phi_s = 0.61$, $N_O = 1.38$, $c = 0.54$, and $J_c = 0.31$ and are based on 5000 replications per condition.

References

Aldrich, J. W., Garcia, H. R., & Mena, G. (1987). Habituation as a necessary condition for maintenance rehearsal. *Journal of Memory and Language, 26*, 632–637.

Anderson, J. A. (1991). Why, having so many neurons, do we have so few thoughts? In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 477–507). Hillsdale, NJ: Erlbaum.

Anderson, J. A. (1995). *An introduction to neural networks*. Cambridge, MA: MIT Press.

Anderson, J. A., Silverstein, J. W., Ritz, S. A., & Jones, R. S. (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. *Psychological Review, 84*, 413–451.

Baddeley, A. D. (1986). *Working memory*. Oxford: Oxford University Press.

Baddeley, A. D., & Lewis, V. (1984). When does rapid presentation enhance digit span? *Bulletin of the Psychonomic Society, 22*, 403–405.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior, 14*, 575–589.

Bakeman, R., & McArthur, D. (1996). Picturing repeated measures: Comments on Loftus, Morrison, and others. *Behavior Research Methods, Instruments, & Computers, 28*, 584–589.

Bamber, D. (1979). State-trace analysis: A method of testing simple theories of causation. *Journal of Mathematical Psychology, 19*, 137–181.

Barrouillet, P., Bernardin, S., & Camos, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology: General, 133*, 83–100.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Brown, G. D. A., Neath, I., & Chater, N. (2007). A ratio model of scale-invariant memory and identification. *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.

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

Chechile, R. A. (1987). Trace susceptibility theory. *Journal of Experimental Psychology: General, 116*, 203–222.

Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology, 33A*, 497–505.

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

Cowan, N., & AuBuchon, A. M. (2008). Short-term memory loss over time without retroactive stimulus interference. *Psychonomic Bulletin & Review, 15*, 230–235.

Cowan, N., Elliott, E. M., Saults, J., Nugent, L. D., Bomb, P., & Hismjatullina, A. (2006). Rethinking speed theories of cognitive development: Increasing the rate of recall without affecting accuracy. *Psychological Science, 17*, 67–73.

Cowan, N., Nugent, L. D., Elliott, E. M., & Geer, T. (2000). Memory-search and rehearsal processes and the word length effect in immediate recall: A synthesis in reply to Service. *Quarterly Journal of Experimental Psychology, 53A*, 666–670.

Cowan, N., Saults, J. S., Elliott, E. M., & Moreno, M. V. (2002). Deconfounding serial recall. *Journal of Memory and Language, 46*, 153–177.

Cowan, N., Wood, N. L., Nugent, L. D., & Treisman, M. (1997). There are two word-length effects in verbal short-term memory: Opposed effects of duration and complexity. *Psychological Science, 8*, 290–295.

Farrell, S. (2006). Mixed-list phonological similarity effects in delayed serial recall. *Journal of Memory and Language, 55*, 587–600.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review, 9*, 59–79.

Farrell, S., & Lewandowsky, S. (2003). Dissimilar items benefit from phonological similarity in serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 838–849.

Fitzgerald, P., & Broadbent, D. E. (1985). Order of report and the structure of temporary memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 217–228.

Gavens, N., & Barrouillet, P. (2004). Delays of retention, processing efficiency, and attentional resources in working memory span development. *Journal of Memory and Language, 51*, 644–657.

Glover, S., & Dixon, P. (2004). Likelihood ratios: A simple and flexible statistic for empirical psychologists. *Psychonomic Bulletin & Review, 11*, 791–806.

Guérard, K., & Tremblay, S. (in press). Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Hanley, J. A., Negassa, A., Edwardes, M. D. d., & Forrester, J. E. (2003). Statistical analysis of correlated data using generalized estimating equations: An orientation. *American Journal of Epidemiology, 157*, 364–375.

Højsgaard, S., Halekoh, U., & Yan, J. (2006). The R package geepack for generalized estimating equations. *Journal of Statistical Software*, 15(2). Available at: <http://www.jstatsoft.org> (accessed 10 April, 2008).

Hudjetz, A., & Oberauer, K. (2007). The effects of processing time and processing rate on forgetting in working memory: Testing four models of the complex span paradigm. *Memory & Cognition, 35*, 1675–1684.

Jolicoeur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology, 36*, 138–202.

Keppel, G. (1991). *Design and analysis: A researcher's handbook*. Englewood Cliffs, NJ: Prentice-Hall.

Knoedler, A. J., Hellwig, K. A., & Neath, I. (1999). The shift from recency to primacy with increasing delay. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 474–487.

Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

Lewandowsky, S., & Farrell, S. (in press). Short-term memory: New data and a model. *The Psychology of Learning and Motivation*.

Lewandowsky, S., & Oberauer, K. (in press). The word length effect provides no evidence for decay in short-term memory. *Psychonomic Bulletin & Review*.

Lewandowsky, S., Brown, G. D. A., Wright, T., & Nimmo, L. M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short-term memory for serial order. *Journal of Memory and Language, 54*, 20–38.

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., & Farrell, S. (2008). Phonological similarity in serial recall: Constraints on theories of memory. *Journal of Memory and Language, 58*, 429–448.

Loftus, G. R., Oberg, M. A., & Dillon, A. M. (2004). Linear theory, dimensional theory, and the face-inversion effect. *Psychological Review, 111*, 835–863.

Naveh-Benjamin, M., & Jonides, J. (1984). Maintenance rehearsal: A two-component analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 369–385.

Neath, I. (1993). Distinctiveness and serial position effects in recognition. *Memory & Cognition, 21*, 689–698.

Nimmo, L. M., & Lewandowsky, S. (2005). From briefs gaps to very long pauses: Temporal isolation does not benefit serial recall. *Psychonomic Bulletin & Review, 12*, 999–1004.

Oberauer, K. (2003). Understanding serial position curves in short-term recognition and recall. *Journal of Memory and Language, 49*, 469–483.

Oberauer, K., & Lewandowsky, S. (in press). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review*.

Page, M. P. A. (2006). *Time and decay in the immediate serial recall of verbal material*. Paper presented at the international conference on memory, Sydney, Australia.

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105*, 761–781.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220–244.

Pashler, H., Johnston, J. C., & Ruthruff, E. (2000). Attention and performance. *Annual Review of Psychology, 52*, 629–651.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*, 437–442.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58*, 193–198.

Phaf, R. H., & Wolters, G. (1993). Attentional shifts in maintenance rehearsal. *American Journal of Psychology, 106*, 353–382.

Posner, M. I., & Rossman, E. (1965). Effect of size and location of informational transforms upon short-term retention. *Journal of Experimental Psychology, 70*, 496–505.

Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J., & Johnson, M. R. (2007). Refreshing: A minimal executive function. *Cortex, 43*, 135–145.

Rickard, T. C., & Bajic, D. (2004). Memory retrieval given two independent cues: cue selection or parallel access? *Cognitive Psychology, 48*, 243–294.

Ruthruff, E., Pashler, H. E., & Klaassen, A. (2001). Processing bottlenecks in dual-task performance: Structural limitation or strategic postponement? *Psychonomic Bulletin & Review, 8*, 73–80.

Schweickert, R., & Boruff, B. (1986). Short-term memory capacity: Magic number or magic spell? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 419–425.

Vallar, G., & Baddeley, A. D. (1982). Short-term forgetting and the articulatory loop. *Quarterly Journal of Experimental Psychology, 34A*, 53–60.

Wilson, M. (1987). MRC psycholinguistic database: Machine usable dictionary (Version 2.00). Avialable at: <http://www.psy.uwa.edu.au/MRCDatabase/uwa_mrc.htm> (accessed 7/7/2007).