Running head: SIMPLE AND COMPLEX SPAN

Turning Simple Span into Complex Span: Time for Decay or Interference from Distractors?

Stephan Lewandowsky, Sonja M. Geiger, Daniel B. Morrell

University of Western Australia

Klaus Oberauer

Universität Zürich

Stephan Lewandowsky

School of Psychology

University of Western Australia

Crawley, W.A. 6009, AUSTRALIA

lewan@psy.uwa.edu.au

URL: http://www.cogsciwa.com

Abstract

We investigated the effects of the duration and type of to-be-articulated distractors during encoding of a verbal list into short-term memory (STM). Distractors and to-be-remembered items alternated during list presentation, similar to the complex-span task that underlies much of working-memory research. According to an interference model of STM, known as SOB (Serial Order in a Box; Farrell & Lewandowsky, 2002), additional repeated articulations of the same word in between list items should cause minimal further disruption of encoding into STM even though the retention interval for early list items is increased. SOB also predicts that the articulation of several different distractor items should lead to much enhanced disruption if the distractor interval is increased. Those predictions were qualitatively confirmed in four experiments which found that it is the type of distractors, not their total duration, that determines the success of encoding a list into STM. The results pose a challenge to temporal models of complex-span performance, such as the Time-Based Resource Sharing model (TBRS; Barrouillet, Bernardin, & Camos, 2004). The results add to a growing body of evidence that memory for the short term is not exclusively governed by purely temporal processes.

Turning Simple Span into Complex Span: Time for Decay or Interference from Distractors?

There has been much recent interest in the role of time in short-term memory (STM). Perhaps like no other psychological process, memory is inextricably linked to the passage of time—neither successful retention nor forgetting can be observed without reference to a temporally preceding episode. Nonetheless, however obvious the involvement of time, its role need not be *causal*, and theoreticians are divided as to whether or not time is causally related to memory performance (e.g., Brown, Neath, & Chater, 2007; Burgess & Hitch, 1999; vs. Lewandowsky & Farrell, 2008b; Nairne, 1990). Theoreticians have differentiated between at least two notions of time; namely *relative* time, which refers to the spacing between several events relative to a later vantage point, and *absolute* time, which refers to elapsed time between a past event and the present. Here, we are primarily concerned with absolute time, and in particular its contribution to forgetting during on-going encoding of a verbal list for immediate forward serial recall. Nonetheless, to delineate the empirical and theoretical playing field, we begin with a brief discussion of relative time.

Time and Short-Term Memory

According to the venerable temporal-distinctiveness notion (e.g., Brown et al., 2007), events are remembered better if they are temporally more "distinct." Distinctiveness relies on relative time: For example, a list item that was presented 5 s ago will be more distinct if it was separated from its preceding neighbor by 10 s rather than by 1 s. That is, even though the critical item has been retained for an equal amount of time (5 s) in both cases, its temporal location *relative* to another event is assumed to determine performance. Specifically, distinctiveness theories postulate that the temporally-crowded case (1 s gap) leads to worse recall than the temporally-isolated case (10 s gap). Despite initial suggestions that temporal isolation has a beneficial effect on short-term memory (e.g., Neath & Crowder, 1996), much recent evidence shows that serial retrieval from short-term memory is largely immune to the effects of temporal separation (e.g., Lewandowsky, Brown, Wright, & Nimmo, 2006; Nimmo & Lewandowsky, 2005, 2006). Specifically, it is now known that when list items are separated by unpredictably varying intervals, and when encoding strategies such as subjective grouping are adequately controlled, temporal isolation does not facilitate immediate serial recall. That is, contrary to the expectations of temporal distinctiveness, the lists A....B....C.D and A.B.C.D (where each "." represents a unit of time) engender equal recall of item B (e.g., Lewandowsky et al., 2006). In a related vein, Oberauer and Lewandowsky (2008) showed that when *all* study items were separated by longer temporal intervals—thus considerably delaying retrieval of early list items because list presentation was slowed—performance remained largely unchanged. Overall, there is now a fairly good consensus that temporal isolation plays at most a minor role in forward serial recall (see Geiger & Lewandowsky, 2008, for a summary of the evidence. There are important exceptions involving unconstrained report order, e.g., Brown, Morin, & Lewandowsky, 2006; Lewandowsky, Nimmo, & Brown, 2008; Lewandowsky, Brown, & Thomas, 2009, and unpredictable list lengths, Geiger & Lewandowsky, 2008, but because those do not involve standard serial recall they are not relevant here.)

Turning to the role of absolute time, temporally-based theories predict that the passage of time will necessarily lead to reduced memory performance. This prediction is most readily understood within the context of decay models, which postulate that an item receives some memory strength upon being encoded, which then inexorably declines over time (e.g., Page & Norris, 1998). Time-based forgetting is also predicted by distinctiveness theories if relative time is controlled: All other things being equal, an item will become less distinct as it recedes into the temporally crowded past, and distinctiveness models thus also expect performance to decline with (absolute) time (provided relative time is constant; for details, see Lewandowsky, Duncan, & Brown, 2004). Recent examinations of the role of absolute time have, however, been largely negative. For example, Lewandowsky et al. (2004) asked participants to recall a short list of letters in forward order at various speeds, in synchrony with an imaginary metronome involving either a slow, medium, or fast beat. Notwithstanding the considerable increase in retention time when the retrieval pace was slowed, recall performance was unaffected by the passage of time. This result has been replicated and extended by Cowan et al. (2006) and by Oberauer and Lewandowsky (2008).

It thus appears that, at least with forward serial recall, there is little evidence for a causal role of time in short-term memory (see Lewandowsky, Oberauer, & Brown, 2009, for a recent review). This conclusion appears warranted in particular because in all of the studies just mentioned, a potential compensatory contribution of rehearsal was experimentally eliminated. Compensatory rehearsal presents a core difficulty in the study of time and memory: Whenever forgetting is found to be absent, decay theorists can claim that people rehearsed the list during the retention period, thereby masking any effect of decay. To control for this possibility, Lewandowsky et al. (2004) asked people to repeat a single irrelevant word ("distractor" from here on) aloud in between retrievals. There is broad consensus that this "articulatory suppression" (AS from here on) blocks rehearsal (Baddeley, 1986, p. 37, p. 86; Baddeley & Lewis, 1984; Page & Norris, 1998, p. 764, p. 770), thus providing an opportunity for decay to express itself if it exists. Taking efforts to suppress rehearsal a step further, Oberauer and Lewandowsky (2008) additionally asked people to perform a symbolic choice task together with AS in the experimentally varied delay periods, and despite the addition of this second distractor task no time-based forgetting was observed. Likewise, the absence of temporal isolation effects mentioned earlier (Lewandowsky et al., 2006; Nimmo & Lewandowsky, 2005, 2006) was observed while people engaged in AS during encoding, and the lack of an overall temporal separation effect reported by Oberauer and Lewandowsky (2008) also arose in the presence of AS. It follows that the pervasive absence of temporal effects in short-term serial recall can hardly reflect a coincidental equilibrium between the two opposing processes of decay and rehearsal.

Although these results converge on a fairly straightforward conclusion, they raise at least one puzzle: If forgetting is not time-based, what, then, causes forgetting? Lewandowsky, Geiger, and Oberauer (2008) provided a partial empirical answer to this question by testing the predictions of a non-temporal model, known as SOB (Serial Order in a Box; Farrell & Lewandowsky, 2002; Lewandowsky & Farrell, 2008b), which explains forgetting on the basis of interference. According to SOB, the articulation of distractors causes their obligatory encoding into memory (Oberauer & Lewandowsky, 2008). A fundamental principle of SOB is that encoding is governed by novelty; information that is novel is encoded strongly whereas information that resembles the existing contents of memory is encoded with reduced strength. It follows that subsequent to the first encoding of a distractor, its additional repetitions will have little novelty and hence will not be encoded with great strength. In consequence, there is minimal forgetting because memory for the list items is not disrupted by subsequent encoding, enabling SOB quantitatively to account for data that showed no time-based forgetting from STM (Oberauer & Lewandowsky, 2008). An obvious corollary of novelty-sensitive encoding is that if the identity of successive distractors were to change, their greater novelty should increase their interfering effect and more forgetting should be observed—this is exactly what Lewandowsky, Geiger, and Oberauer (2008) found in several experiments. Repeated articulation of the same distractor (e.g., “candle”) in between retrievals caused no additional forgetting compared to a single articulation of that distractor in between two recalls. By contrast, articulation of different words (e.g., “candle”, “frog”, “chair”) caused considerable additional forgetting compared to repetitions of “candle” even though retention time was extended by a virtually equal amount. These results confirm the predictions of SOB and demonstrate that whether or not forgetting of a just-encoded list increases with longer delays depends on the nature of the interfering material processed during that delay—specifically, its novelty vis-a-vis earlier distractor material—rather than the passage of time per se.

Time and Working Memory

The results just summarized, showing that neither absolute nor relative time plays an important role in immediate serial recall, appear to be at odds with prevailing developments in working-memory (WM) research that appeal to temporal processes. Both WM and STM refer to retention over the short term and are therefore closely related concepts. However, whereas STM tasks only involve storage and immediate serial recall of a list, known as simple-span tasks, examinations of WM typically involve some additional cognitive processes beyond encoding and retrieval (Unsworth & Engle, 2007). For example, in the “complex-span” task, study items may alternate with, for instance, mental arithmetic problems (Turner & Engle, 1989). People might be presented with a sequence such as $2 + 3 = 5?$, A, $5 + 1 = 7?$, B, $\ldots$, where the equations have to be judged for correctness and the letters must be memorized for immediate serial recall after the sequence has been completed. The nature of the processing task turns out to matter little; the effects of arithmetic, simple reading of digits, and mere articulation of distractors are qualitatively identical (e.g., Barrouillet et al., 2004).

The addition of a processing task has profound implications: First, performance on complex-span tasks is lower than on simple-span tasks, implying that the processing task in between list items causes forgetting of the earlier, already-encoded items (or alternatively, interferes with encoding of the subsequent items). Second, complex-span performance accounts for a whopping half of the variance among individuals in measures of general fluid abilities (Kane, Hambrick, & Conway, 2005), which is considerably more than the variance explained by simple-span tasks (e.g., Conway, Cowan, Bunting, Therriault, & Minkoff, 2002; Engle, Tuholski, Laughlin, & Conway, 1999). The strong link between WM performance and fluid intelligence highlights the importance of understanding the reasons for short-term forgetting—it is forgetting during distracting processes in between encoding of a list, after all, that limits performance in complex-span tasks.

In the WM literature, forgetting in the complex-span task is frequently (though not always; Saito & Miyake, 2004) explained by temporal decay. For example, according to the Time-Based Resource Sharing model (TBRS; Barrouillet et al., 2004), memory traces rapidly fade while people attend to the processing task (e.g., arithmetic). Unlike earlier decay-based models of the complex-span task (Towse, Hitch, & Hutton, 2000), the TBRS postulates a rapid attentional mechanism that can be deployed to "refresh" memory traces. Thus, although memory is thought to decay during processing—because attention can only be deployed to one task at a time—the TBRS suggests that even brief pauses in between processing steps are sufficient for attentional refreshing of memory traces. In consequence, forgetting is thought to be a function not of the absolute processing duration, but of the *proportion* of the time in between two study items during which the attentional mechanism is actually devoted to the processing task. This proportion of time, known as "cognitive load," is a crucial determinant of performance: the lower the cognitive load, the more time there is for refreshing and hence the better performance should be. Conversely, the greater the load, the worse performance is expected to be. Barrouillet and colleagues have presented an impressive array of evidence in support of the role of cognitive load (e.g., Barrouillet et al., 2004; Barrouillet, Bernardin, Portrat, Vergauwe, & Camos, 2007): Increasing the number of distractors while keeping total available processing time constant–thus increasing cognitive load–reduces performance considerably. Likewise, holding the number of distractors constant and shortening the total available time impairs memory. Both manipulations of cognitive load affect the restoration time available for refreshing in between processing of individual distractors. Therefore, there is now little doubt that brief pauses in between processing steps permit some restorative process to operate, with attendant beneficial effects on memory.

It is less clear, though, whether forgetting is caused by decay, so that longer time taken up by processing leads to more forgetting when restoration time is held constant. Researchers from the TBRS team have presented results showing that extending processing duration while keeping restoration time constant reduces memory performance (Portrat, Barrouillet, & Camos, 2008); however, we have argued elsewhere that these data are inconclusive (Lewandowsky & Oberauer, in press). Moreover, Oberauer and Lewandowsky (2008) have shown that adding further distractors while cognitive load was maximal (because pauses between distractors were eliminated) has no effect on performance. Thus, despite the impressive empirical success of the TBRS model, there remain doubts about one of its core assumption, namely, that forgetting in WM is caused by time-based decay.

This article investigates the role of time and interference as potential causes of forgetting during a distracting processing task in between study items in a complex span paradigm by testing quantitative predictions of the SOB theory (Farrell & Lewandowsky, 2002; Lewandowsky & Farrell, 2008b) and contrasting them to the predictions of the TBRS (Barrouillet et al., 2004). To foreshadow, we begin by presenting SOB's threefold predictions: a single intra-list distractor should disrupt performance but the extent of that disruption should not increase with additional repetitions of the same distractor. By contrast, when several different distractors are articulated together in between list items (in what we call a "burst" from here on), the extent of disruption should increase with the number of distractors.[1] We next present the competing predictions of the TBRS, which are straightforward: If cognitive load is nearly or exactly 1 (i.e., all available time in between list items is taken up by processing of distractors), then adding additional distractors must lead to a decrement in performance, irrespective of whether or not the same word is repeated or different words are articulated. We then report four experiments whose consistent results are in accord with SOB's predictions: Extending intra-list intervals by increasing the number of distractors in each burst impaired recall only if distractors within a burst differed from each other, whereas increasing burst size had no effect when the same distractor was repeated multiple times. The data strongly challenge the TBRS.

SOB Predictions

Farrell and Lewandowsky (2002) presented a distributed model of short-term serial recall called SOB, for "Serial-Order in a Box," a name that acknowledged the model's reliance on the Brain-State-in-a-Box algorithm (Anderson, Silverstein, Ritz, & Jones, 1977). SOB assumes that items are represented by vectors of features that are encoded into memory by first being associated to a positional marker, before being superimposed onto a common weight matrix. All encoding and retrieval processes in SOB are entirely dependent on events—such as encoding or retrieval of an item—and are not affected by the passage of time. Thus, whatever forgetting occurs in SOB results from interference not temporal decay.

One core property of SOB is that all 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. 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 core 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, 2008b), a prediction that has been confirmed repeatedly (Farrell, 2006; Farrell & Lewandowsky, 2003). Finally, and most important in the present context, novelty-sensitive encoding has also enabled SOB to handle the finding that repeating the same distractor several times leads to little or no further forgetting (Lewandowsky et al., 2004), while at the same time accounting for the additional forgetting that is observed when the nature of the distractors is changed (Lewandowsky, Geiger, & Oberauer, 2008).

For the present studies, we obtained predictions from SOB using published parameter settings (fits to Experiment 1 in Oberauer & Lewandowsky, 2008). Details of the simulation are presented in the Appendix. We derived predictions for four distinct relationships among distractors by orthogonally varying the extent of *between-burst variation* and the extent of *within-burst variation.* We use the term *burst* to refer to a series of one or more distractor(s) to be articulated in between two list items. Between-burst variation refers to whether or not the same set of distractor(s) was articulated after every item. Within-burst similarity refers to whether each burst involved the same or different items. This latter variable becomes relevant only when bursts contain more than one distractor.

Figure 1 summarizes SOB's predictions for lists on which each to-be-remembered item (bar the last one) was followed by distractor activity. The three lines in each of the 4 panels represent conditions in which zero (baseline), one, or three distractors follow each item. Rows of panels correspond to the two levels of within-burst variation. The top two panels show predictions for bursts involving repetition of the same distractor (e.g., "office", "office", "office"). We refer to these bursts as *simple* from here on. The bottom panels show predictions for bursts consisting of different words (if a burst involved more than one word). For example, people might articulate the words "office", "question", and "yearly" after a study item. We refer to these bursts as *complex* from here on.

The left- and right-hand panels in Figure 1 compare predictions for the two levels of between-burst variation. The left column of panels shows the predictions when the same bursts—irrespective of whether they are simple or complex—are repeated after each study item. We refer to this level of between-burst variation as *steady* from here on. The right-hand panels, by contrast, show predictions when the identity of bursts—irrespective of whether they are simple or complex—changed across list positions, a type of between-burst variation that we call *changing* from here on.

To summarize, ignoring the baseline condition, the top-left panel corresponds to people saying the same word either once or three times after every list item; this corresponds to the experimental manipulation reported by Oberauer and Lewandowsky (2008). The remaining three panels have not been experimentally examined to date and correspond to the following input regime: the top-right panel represents a word being articulated either once or three times after each list item (bar the last one), with the identity of that word changing across serial positions (but not within a burst). Finally, the three-distractor conditions in the bottom panels correspond to people articulating three different words in each burst, with the identity of those three words either repeating across serial positions (left-hand panel; e.g., "office", "question", and "yearly" in each burst) or changing continuously (right-hand panel; e.g., "office", "question", and "yearly" after the first item followed by "quokka", "marron", "bilby" after the second list item, and so on).[2]

The predictions of SOB are readily summarized. (a) Presentation of a single distractor during list encoding causes a dramatic impairment of performance relative to the quiet condition across all serial positions. (b) Additional repetitions of the same distractor in a simple burst (top row) does not cause any further performance decrement. (c) If a burst involves articulation of different items, additional distractors cause an additional performance decrement (bottom row). (d) To a first approximation, the preceding three effects of distractors at encoding are fairly evenly spread across all serial positions. (e) There is an effect of between-burst variation: Changing bursts (right-hand panels) generally cause more interference than steady bursts (left-hand panel). (f) The effect of between-burst variation is attenuated for complex bursts of three distractors, where it does not matter whether the three words within each burst change across serial positions or are repeated. By implication, the overall effect of the number of distractors is attenuated for changing bursts.

These predictions arise directly from the core principles of SOB: Owing to novelty-sensitive encoding, only the first distractor within a simple burst is encoded with full strength, whereas its subsequent repetitions receive little further encoding. In consequence, increasing the number of distractors has little or no effect with simple bursts. Owing to the same principle of novelty-sensitive encoding, all distractors within a complex burst, by contrast, are equally novel and will thus be encoded with full strength. In consequence, with complex bursts, more distractors will cause additional forgetting. The effect of between-burst variation arises in the same way. Between-burst variation has a smaller effect than within-burst variation, because although adjacent context markers partially overlap they nonetheless differ from each other: As a consequence, distractors that are repeated in different bursts after different list items are regarded as more novel than distractors that are repeated within the same burst.

TBRS Predictions

Unlike plain decay models, the TBRS postulates two opposing processes—decay and attentional refreshing—whose balance determines performance in complex-span tasks. These two opposing processes imbue the model with considerable explanatory power, but they also render its test more challenging. One way in which the model can be tested is by experimentally eliminating the presumed refreshing process by increasing cognitive load to 1. When cognitive load is 1, by definition all available time in between items is taken up with processing and no time remains for refreshing—it follows that extending the total duration of processing must lead to additional forgetting.

These predictions are illustrated in Figure 2. The top two panels (A and B) reiterate that the addition of extra processing steps or distractors need not have an effect when cognitive load is safely below 1: The time available for restoration (upward triangles labeled R in the figure) fully compensates for the decay that occurs during processing (downward triangles P); hence, adding further distractors has no effect on memory strength. This prediction was experimentally confirmed by Barrouillet et al. (2004, Experiment 3).

However, as cognitive load increases, a point will be reached at which the refreshing possible within two processing steps can no longer fully compensate the decay occurring during a processing step. From this point on, each cycle of a processing step, followed by some free time for refreshing, will engender a net loss of memory strength. Increasing cognitive load further eventually eliminates refreshing altogether. The bottom two panels (C and D) show the model's predictions, hitherto untested, when cognitive load is 1. No restoration can be performed because there is no free time available for attentional refreshing (triangles R are absent), and hence the addition of further distractors inexorably leads to further decay (compare final memory strength in panel D to final strength in panel C). Thus, whether the TBRS predicts more forgetting with more distractors depends on the level of cognitive load - with a high load, more distractors must lead to more forgetting. This prediction is independent of the nature of the distractors. The TBRS assumes no role for interference and it is only the duration of processing that matters (Barrouillet et al., 2007, p. 572). Therefore, at cognitive load near 1, the TBRS predicts that bursts with more distractors lead to more forgetting than bursts with fewer distractors, and that neither within-burst nor between-burst similarity make any difference.

The present experiments were designed to adjudicate between the contrasting predictions of the TBRS (panels C and D in Figure 2) and of SOB (Figure 1). It is important to note that the predictions of TBRS can only be tested when two conditions are met: First, the processing of distractors must occupy the attentional bottleneck (Barrouillet et al., 2004). Barrouillet et al. (2007) showed that tasks that fail to occupy the attentional bottleneck for the duration of processing—such as simple detection tasks that do not involve a decision between response alternatives—do not block attentional refreshing. Second, cognitive load must be verifiably high enough to prevent refreshing from fully compensating the effect of decay. Both conditions were met by the experiments that follow; we present supporting evidence in the General Discussion, after all data have been presented.

Experiment 1

The first experiment included three orthogonal experimental variables: Within-burst variation (simple vs. complex), the number of distractors within a burst (1 vs. 3), and between-burst variation (steady vs. changing). The first variable was manipulated between subjects and the remainder within subjects. In addition, the design included a quiet baseline condition involving lists that were presented without articulation of distractors.

Method

Participants. Participants were 40 members of the University of Western Australia campus community. They participated voluntarily in a 1-hour session for which they were remunerated with A$10 ($N = 25$) or were awarded credit towards a course requirement ($N = 15$). The participants' ages ranged from 16 to 42. Participants were randomly assigned to the two experimental groups involving either simple ($N = 21$) or complex ($N = 19$) distractors.

Apparatus and materials. A Windows computer running a Matlab program designed using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) was used to present stimuli and for the recording of responses.

Participants had to encode and orally recall lists of 5 consonants in their correct forward order. List items were sampled randomly and without replacement from a pool consisting of all consonants except Q and Y.

The distractor words consisted of nouns (written frequency < 89 according to Kucera & Francis, 1967) with 2 or 3 syllables and 6-14 letters in length, taken from the MRC psycholinguistics database (Coltheart, 1981; Wilson, 1988). Distractors were drawn from a pool of 370 words, thus ensuring that every distractor for a given participant was unique and not repeated elsewhere in the experiment.

Procedure. Each trial commenced with a cue that specified the number and type of distractors for that trial (e.g. "1 Changing" or "3 Steady"). The cue appeared in the center of the screen for two seconds. Following this, a fixation cross was centrally presented for 0.9 seconds, followed by presentation of the first list item. List items were presented centrally for 900 ms and had to be read aloud. Items were either followed by presentation of the next list item for articulation, or by a list of one or more distractors that also had to be read aloud. Distractors appeared rapidly, one after the other (200 ms per distractor) on separate rows at the top of the screen. Once presented, each distractor remained visible until all distractors had been articulated.

Participants articulated either zero, one, or three distractors after each list item with the exception of the last. The identity of distractor words after each item determined the level of between-burst variation for a given trial: The distractors either changed (e.g., "office" after the first item followed by "question" after the second, etc) or were steady (e.g. ,"office" after each of the first 4 to-be-remembered items). Within-burst variation was likewise instantiated by the identity of the distractors when bursts contained more than one item. Participants in the simple condition articulated the same word three times (e.g., "office, office, office") whereas participants in the complex condition articulated three different words (e.g., "office, question, yearly"). Note that when list items were separated by only a single distractor, simple and complex bursts are identical.

Following the participant's articulation of each list item and its associated distractors, the experimenter presented the next list item (and distractors, if any) by pressing the space bar. The experimenter ensured that people articulated distractors and list items continuously without permitting any breaks that could have been used for refreshing of memory items.

After presentation and articulation of the last list item, a question mark appeared in the center of the screen. This prompted participants to recall the list items orally in the order in which they were presented. The experimenter entered each response on the keyboard. After the experimenter entered the last list item, there was a two second break after which the cue for the next trial appeared in the center of the screen.

Participants were instructed to recall all 5 items to the best of their ability. Participants were encouraged to guess if they were unsure of an item, but they could say "pass" as a last resort to indicate an omission.

The experimental trials were preceded by 5 practice trials (one list of each type): quiet baseline, single-distractor steady, three-distractors steady, single-distractor changing, three-distractors changing. Participants then undertook 60 experimental trials that included 12 of each of the above types of list. Trials were randomized without constraint for each participant and participants were given a self-paced break after every fifteen trials.

Results

Individual differences. Individual correct-in-position recall performance (averaged across all conditions for each participant) ranged from .38 to .83 across individuals, safely away from both ceiling and floor. All participants were retained for analysis.

Timing manipulation. To ensure that the timing manipulation was successful, the mean times (across serial positions) between pairs of to-be-remembered items were compared between lists involving one and three distractors, respectively. Times were measured from a study item's offset, through the participant's reading of distractor items and finishing as the experimenter pressed a key to present the next list item (in the baseline condition these times were not meaningful and hence that condition was excluded from this analysis). Table 1 shows the means and standard deviations of the encoding times.

A 3-factorial between-within ANOVA with Within-burst variation (2) $\times$ Between-burst variation (2) $\times$ Number of distractors (2) as factors was conducted on the articulation times. The ANOVA revealed large main effects of number of distractors (1 vs. 3), $F(1, 38) = 1002.51, MSE = .036, p < .0001$, partial $\eta^2 = .98$, and between-burst variation (steady vs. changing), $F(1, 38) = 74.21, MSE = .021, p < .0001$, partial $\eta^2 = .66$, but not within-burst variation (simple vs. complex), $F(1, 38) < 1$. In absolute terms, the main effect of between-burst variation was small (200 ms difference between steady and changing) compared to the magnitude of the effect caused by increasing the number of distractors (950 ms).

The analysis additionally uncovered two-way interactions between all three variables; that is, Number of distractors × Within-burst variation, $F(1, 38) = 20.46$, $MSE = .036, p < .0001$, partial $\eta^2 = .35$; Number of distractors × Between-burst variation, $F(1, 38) = 12.12$, $MSE = 0.008, p < .005$, partial $\eta^2 = 0.23$; and Within-burst variation × Between-burst variation, $F(1, 38) = 10.21$, $MSE = .021, p < .005$, partial η^2 $= .21$. These interactions reflected the fact that additional distractors had a larger effect with complex bursts compared to simple bursts (and likewise with changing vs. steady bursts), and that the effect of within-burst variation was greater with changing than with steady bursts. The over-arching three-way interaction between all variables was also significant, $F(1, 38) = 12.34$, $MSE = .008, p < .005$, partial $\eta^2 = 0.23$. This interaction reflected the fact that adding two more distractor words had a particularly large effect with complex bursts when those were changing (compare the bottom right-hand cell in Table 1 to the other table entries).

Overall, we conclude that the timing manipulation in Experiment 1 was successful, and that increasing the number of distractors between items at encoding successfully increased the temporal spacing between list items. We defer further consideration of the impact of the timing manipulation on recall until after all the data have been presented.

Recall performance. We first compared the general performance level of the two groups by examining performance on the quiet baseline lists. A Group (2) × Serial position (5) between-within ANOVA involving only the correct-in-position scores for the baseline lists yielded the expected highly significant effect of serial position, $F(4, 152) = 25.36$, $MSE = .007$, $p < .0001$, partial $\eta^2 = .40$, but failed to show a main effect of group, $F(1, 38) = 1.27$, $MSE = .034$, $p > .10$, partial $\eta^2 = .03$. Likewise, the interaction between both variables was non-significant, $F(4, 152) = 2.01$,

$MSE = .007, .05 < p < .10$, partial $\eta^2 = 0.05$. We conclude that the two groups were comparable in recall performance when no articulation was required during encoding ($M = .90$ and $M = .88$ for the groups who received simple and complex bursts, respectively). Because the baseline conditions were not fully crossed with the two experimental variables that determined the nature of the remaining 4 list types, the baseline lists were not considered any further.

Figure 3 shows the serial position curves for all conditions using the same layout as for the earlier SOB predictions. The quiet baseline lists are included in all panels for comparison. 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).

The figure suggests that the effect of within-burst variation predicted by SOB was obtained, with the effect of additional distractors (lines labeled 1E vs. 3E) being somewhat larger for complex bursts (bottom row of panels) than for simple bursts (top row). In confirmation of another prediction by SOB, performance with steady bursts (left-hand column of panels) was considerably better than with changing bursts (right-hand column).

Because the effect of within-burst variation was of principal theoretical interest but perhaps small in magnitude, we explored it further by averaging across the between-burst variation variable. The two left-hand panels in Figure 4 present these averaged data, highlighting the small effect of additional distractors with simple bursts (top-left panel) and contrasting it with the considerably larger effect for complex bursts (bottom-left). The larger panel on the right shows the difference in performance between lists with a single distractor and those with three distractors for simple and complex bursts simultaneously. The panel shows that with the exception of the first serial position, the number of distractors in a simple burst mattered little. The number of distractors in a complex burst, by contrast, mattered considerably more, except for the last list item which was not followed by a burst.

Statistical confirmation of the pattern in the figures was provided by a 2 (Within-burst variation) × 2 (Between-burst variation) × 2 (Number of distractors) × 5 (Serial position) between-within subjects ANOVA. The results of the analysis are summarized in Table 2.

The ANOVA yielded a number of effects, most of which were expected either on the basis of conventional wisdom (e.g., the main effect of serial position) or on the basis of SOB's predictions (i.e., the interactions between number of distractors and the two distractor-variation variables, as described above). The only partially unexpected effect involved the interaction of number of distractors(ND) with serial position (SP). This interaction was not predicted by SOB but also comes as little surprise because the current experiment did not involve distractors after the last list item (unlike the simulations generating the SOB predictions shown in Figure 1. We return to this issue in the context of Experiment 3 which included distractors after the terminal item.)

The two theoretically most interesting effects are the interactions of number of distractors (ND) with between-burst variation (BB) and with within-burst variation (WB): The former interaction is clearly apparent from Figure 3 and the latter is particularly visible in Figure 4. Both interactions reflect the fact that the disruptive effect of adding further distractors was roughly doubled by the burst manipulations: For between-burst variation, the effect grew from 6% for changing bursts ($M = .48$ and $M = .42$ for 1 and 3 distractors, respectively) to 10% for steady bursts ($M = .61$ vs. $M = .51$). For within-burst variation, the additional disruption was 5% for simple bursts ($M = .54$ and $M = .49$ for 1 and 3 distractors, respectively) but 11% for complex bursts ($M = .55$ vs. $M = .44$). Both interactions were in the direction predicted by SOB, and the remainder of this article will primarily focus on the latter.

Discussion

The results of the first study were at least qualitatively consistent with the predictions of SOB: Complex bursts had a larger disruptive effect on memory than simple bursts, notwithstanding the fact that both delayed encoding of the next list item—thus increasing the retention interval for the preceding items—by a roughly equal amount. Likewise, the disruptive effect of additional distractors was greater when bursts remained the same across serial positions than when they changed; this effect may appear counter-intuitive at first glance until one considers the fact that single distractors supported considerably better performance when they were identical for the whole list ($M = .61$) than when they changed between items ($M = .51$). Hence, the disruptive effect of additional distractors for steady bursts was assessed against a higher baseline and was therefore larger than with changing bursts.

Nonetheless, the data also diverged from the predictions of SOB in potentially important ways. Specifically, even though additional distractor time caused less disruption with simple bursts than with complex bursts, it nonetheless had a detectable effect (5%; this was significant in a separate ANOVA focusing only on simple bursts; $F(1, 20) = 7.68$, $p < .05$, $MSE = .034$, partial $\eta^2 = .28$). This effect was not expected by SOB and it also deviates from the results of Oberauer and Lewandowsky (2008), who repeatedly found that adding repetitions of the same word at encoding did not lead to any greater disruption compared to a single articulation. One possible reason for this discrepancy in outcomes is that the present study included both steady and changing bursts across trials whereas Oberauer and Lewandowsky only used steady bursts. Another possible reason is the fact that unlike in the studies by Oberauer and Lewandowsky, the terminal item in the present study was not followed by any distractors. Another departure between predictions and data involves the role of serial position: Whereas SOB predicted the additional disruption due to complex distractors to be evenly spread across serial positions, the data clearly showed a convergence, such that performance at the last serial position was immune to within-burst similarity (see right-hand panel of Figure 4).

The remaining experiments were designed to probe further into these two issues raised by Experiment 1. Specifically, Experiment 2 used a within-subjects design to compare simple and complex bursts when both were changing. Experiment 3 (and a companion study) used a nearly identical within-subjects design but additionally included distractors after the last list item in order to examine whether this might introduce an effect of within-burst similarity at the last list position.

Experiment 2

Experiment 2 involved 4 conditions that were all manipulated within subjects: The quiet baseline; a single changing distractor; simple-changing bursts of 3 distractors; and complex-changing bursts of 3 distractors. The experiment focused on changing distractors because in Experiment 1 the effect of within-burst variation was smaller for those than with steady distractors, thus rendering it particularly important to examine the replicability of that effect.

Two contrasts were of particular interest: First, the effect of additional repetitions of the same distractor (i.e., the contrast between 1 vs. 3 simple-changing). This effect (albeit small in magnitude) was present in Experiment 1 but was not predicted by SOB. Second, the contrast between simple and complex bursts, which was predicted by SOB and was present in Experiment 1 but requires replication.

Method

Participants. Participants were 24 first-year psychology students of the University of Western Australia. They participated voluntarily in a 1-hour session for partial fulfillment of course requirements. None had participated in Experiment 1.

Materials and procedure. The experiment employed 4 of the 8 conditions tested in Experiment 1: The quiet baseline, lists with a single (changing) distractor, and two lists with bursts of 3 distractors, which were either complex (and changing) or simple (and changing). Conditions were manipulated across trials within subjects.

Because there were only 4 different list types (as opposed to five in Experiment 1), the number of trials per list type was increased to 15, so the sum of trials stayed constant at 60. In all other respects, the material and procedure were identical to those of Experiment 1.

Results and Discussion

Individual differences. Individual recall performance averaged across all conditions ranged from .39 to .80 across individuals, safely away from both ceiling and floor. All participants were retained for analysis.

Timing manipulation. Encoding times between items were measured as in Experiment 1. Averaged across serial positions, the encoding times were $1.40s$ for a single distractor at encoding, $2.18s$ for simple bursts, and $2.61s$ for complex bursts.

The effect of the timing manipulation was explored with a one-way within-subjects ANOVA with Condition (3) as the single factor (omitting the quiet baseline lists for which encoding times were meaningless). The ANOVA revealed an overall significant effect of the distractor manipulation, $F(2, 46) = 400.5$, $MSE = 0.023$, $p < .001$, partial $\eta^2 = .95$. Planned contrasts revealed large differences between the single distractor and 3-distractor simple bursts, $F(1, 23) = 927.5$, $MSE = 0.016$, $p < 0.0001$, $\eta^2 = .98$, and between the single distractor and 3-distractor complex bursts, $F(1, 23) = 609.4$, $MSE = 0.057$, $p < 0.001$, $\eta^2 = .96$, thus attesting to the success of the timing manipulation. The difference between the two long bursts (simple vs. complex) was smaller but still significant, $F(1, 23) = 70.0$, $MSE = 0.062$, $p < 0.001$, $\eta^2 = .75$. Although the complex

3-distractor bursts took .43 seconds longer to articulate than the corresponding simple bursts, this difference was smaller than the mean increase of encoding time from 1 distractor to the 3-distractor conditions (mean increase of 1.0 second).

Recall performance. Figure 5 shows the correct-in-position recall for all four conditions in Experiment 2. A 5 (Serial position) × 4 (Condition) within-subjects ANOVA revealed a significant effect of condition, $F(3, 69) = 88.44$, $MSE = .05$, $p < .0001$, partial $\eta^2 = .79$ and of serial position, $F(4, 92) = 29.96$, $MSE = .04$, $p < .0001$, partial $\eta^2 = .57$, as well as an interaction between both variables, $F(12, 276) = 6.89$, $MSE = .02$, $p < .0001$, partial $\eta^2 = .23$, thus confirming the obvious pattern in the figure.

These effects were further explored by contrasts. For the main effect of condition, the baseline lists were found to differ from the lists with 1 distractor, $F(1, 23) = 118.76$, $MSE = .62$, $p < .0001$, $\eta^2 = .84$, which in turn did not differ from lists with simple bursts of 3 distractors, $F(1, 23) < 1$. This replicates the known detrimental effect of a single distractor compared to a quiet baseline and it also replicates the finding from previous experiments (Oberauer & Lewandowsky, 2008) that adding further repetitions of the same word does not cause a further decrement in performance. Finally, there was a significant difference between simple and complex bursts, $F(1, 23) = 9.70$, $MSE = .28$, $p < .005$, $\eta^2 = .30$, confirming that unlike repetitions of the same word, articulation of several different words generates a significant further performance decrement.

Exploration of the Condition × Serial position interaction by additional contrasts (with serial position coded by its linear or quadratic contrast) revealed that the contrasts involving the three complex distractors both interacted with Serial Position. For the contrast between the single distractors and complex bursts, the interaction with serial position was, $F(1, 23) = 11.1$, $MSE = 0.057$, $p < .01$, partial $\eta^2 = 0.33$. For the contrast between simple and complex bursts of 3 distractors, the interaction with serial position was also significant, $F(1, 23) = 7.97$, $MSE = 0.052$, $p < .01$, partial $\eta^2 = 0.31$, thus confirming the obvious "fanning" in Figure 5. for both of those comparisons. The remaining interaction involving the contrast between single-distractors and simple bursts and serial position was not significant, $F < 1$.

Experiment 2 thus replicated those findings of the first study that were predicted by SOB, showing again that within-burst variation is an important determinant of whether extending distractor time causes additional disruption during encoding into short-term memory. In replication of Oberauer and Lewandowsky (2008) but unlike Experiment 1, extending the duration of simple bursts had no effect in this study. Only when the additional distractors differed from each other was there a large disruptive effect. One limitation of Experiment 2 was that the encoding times differed significantly between simple and complex 3-distractor bursts, with complex bursts taking an additional 430 ms to encode. Although this difference was slight compared to the difference between a single distractor and 3 simple (780 ms) or complex (1.21 s) distractors, it may have contributed to the greater disruption associated with complex bursts. The data nonetheless challenge a purely temporal account because the 780 ms additional time between items on the lists with simple bursts compared to single-distractor lists had no effect despite the fact that recall of the first item was delayed by over 3 seconds (4×780 ms). The remaining experiments addressed this issue by comparing bursts of *four* simple distractors to bursts of three complex distractors, thus equalizing processing times between burst types.

Experiment 3

Experiment 3 introduced three new features: First, and most important, the simple bursts involved 4 repetitions of the same word whereas the complex bursts involved 3 different words. The extra repetition for simple bursts was intended to equalize total time spent on the distractor activity in the two conditions. Second, every list item—including the last one—was now followed by a burst of distractors. This was intended to examine whether within-burst similarity might then have an effect on the last serial position as well. Finally, to maximize the possible contrast between simple and complex bursts, all simple bursts were steady (unlike in Experiment 2) whereas all complex bursts were changing (as in Experiment 2).

Method

Participants and materials. Participants were 24 members of the campus community of the University of Western Australia (14 females, age range 17–59, mean 26.7. Exclusion of two participants who were older than 50 does not alter any of the conclusions) who received A$10 for participation. The distractors were created anew according to the following criteria: All were nouns of between 6-14 letters (2-3 syllables) with a written frequency in excess of 70 (Kucera & Francis, 1967). A total of 325 distractor words were created for this experiment. None had participated in the earlier experiments.

Procedure. The procedure was identical to that of Experiment 2 with the following exceptions: First, the four conditions were (a) quiet baseline with no distractors; (b) lists with one (steady) distractor after each item; (c) lists with simple bursts of 4 repetitions of an item (also steady); and (d) lists with complex changing bursts of 3 items (also steady). Distractors were unique to the trial on which they appeared. Second, in all conditions involving distractors, all list items (including the last one) were followed by distractor(s). Third, there were 12 trials in each condition, for a total of 48 experimental trials.

Results and Discussion

Individual differences. Individual recall performance averaged across all conditions ranged from .30 to .90, with all individuals being within ± 2 standard deviations of the grand mean. All participants were therefore retained for analysis.

Timing manipulation. Averaged across serial positions, the encoding times were 1.15s for the lists with a single distractor, 2.45s for simple bursts of 4 words and 2.42s for complex bursts of 3 words. As intended, adding a fourth repetition to the simple bursts equalized processing times between the two types of bursts.

In confirmation, a one-way within-subjects ANOVA (involving the 3 distractor conditions but not the quiet baseline) yielded a significant main effect of Condition, $F(2, 46) = 642.25$, $MSE = .021$, $p < .0001$, $\eta_p^2 = .97$. Planned contrasts confirmed that the single-distractor condition differed from 3-distractor lists with both simple, $F(1, 46) = 989.11$, $MSE = .021$, $p < .0001$, $\eta_p^2 = .96$, and complex bursts, $F(1, 46) = 936.93$, $MSE = .021$, $p < .0001$, $\eta_p^2 = .95$. Simple and complex bursts did not differ with respect to processing time, $F(1, 46) < 1$.

Recall performance. The serial position curves for all conditions are shown in the left-hand panel of Figure 6. A 5 (Serial position) × 4 (Condition) ANOVA revealed a significant effect of condition, $F(3, 69) = 97.13$, $MSE = .04$, $p < .0001$, $\eta^2 = .81$ and of serial position, $F(4, 92) = 64.60$, $MSE = .02$, $p < .0001$, $\eta^2 = .74$, as well as an interaction between both variables, $F(12, 276) = 4.66$, $MSE = .01$, $p < .0001$, $\eta^2 = .17$.

Follow-up contrasts were conducted to explore the main effect of condition. The baseline lists were found to differ from the lists with a single distractor, $F(1, 23) = 73.48$, $MSE = .62$, $p < .0001$, $\eta^2 = .76$, which in turn did not differ from lists with simple bursts of 4 distractors, $F(1, 23) = 1.60$, $MSE = .27$, $p > .10$, $\eta_p^2 = .07$. In addition, there was a significant difference between simple-steady and complex-changing bursts, $F(1, 23) = 74.39$, $MSE = .16$, $p < .0001$, $\eta^2 = .76$. The pattern of contrasts thus replicated the results of Experiment 2.

To further accentuate the differences between simple and complex bursts, the right-hand panel of Figure 6 plots the net performance deficit associated with additional distractors—that is, 4 vs. 1 and 3 vs. 1—across serial positions. The greater disruption associated with complex bursts is self-evident; note that the time required to process the two types of distractors was virtually identical (differing by only 30 ms per burst). Note also that the mean effect across serial positions of 3 additional repetitions of the same word is close to zero (i.e., the large white circle on the right).

Further contrasts were conducted to explore the interaction between condition and serial position. We highlight two noteworthy results. First, there was a significant interaction involving the linear contrast of serial position and the difference between the two simple-burst conditions (i.e., 1 steady vs. 4 simple-steady), $F(1, 23) = 11.86$, $MSE = .035$, $p < .003$, $\eta_p^2 = .34$. This interaction captured the obvious cross-over between those two serial position curves in the left panel of Figure 6. Second, the contrast between simple-steady and complex-changing bursts did not interact with any of the contrasts of serial position, largest $F(1, 23) = 1.92$, $p > .1$. This reflects the obvious parallelism between the two serial position curves, which differs from the outcome of the first two experiments. We attribute this parallelism to the terminal burst of distractors, which extended the interfering effect of complex bursts to the last serial position.

Generality of results. To extend the generality of the observed effects of within-burst similarity, we conducted a companion experiment ($N = 26$; not reported in detail here) which differed from Experiment 3 only in that participants did not articulate the memoranda. Simple bursts again led to little time-based forgetting whereas complex bursts caused a considerable performance decrement. A 4 (Condition) × 5 (Serial position) within-subjects ANOVA revealed a significant effect of condition, $F(3, 75) = 53.87$, $p < .0001$, $MSE = .054$, serial position, $F(4, 100) = 67.19$, $p < .0001$, $MSE = .021$, and the interaction between both variables, $F(12, 300) = 3.47$, $p < .0005$, $MSE = .011$. Follow-up contrasts confirmed that the baseline list differed from lists with a single distractor, $F(1, 25) = 25.77$, $p < .0001$, $MSE = .68$, which in turn did not differ from lists with 4 simple distractors, $F(1, 25) < 1$. Simple-steady and complex-changing bursts differed considerably from each other, $F(1, 25) = 40.82$, $p < .0001$, $MSE = .53$. Moreover, as in Experiment 3, the contrast between simple-steady and complex-changing bursts did not interact with any of the contrasts (i.e., linear, quadratic, etc.) of serial position, largest $F(1, 25) < 1$. The pattern of contrasts in this companion study thus exactly replicated the results of Experiment 3, including the effect of within-burst similarity for the last serial position.

General Discussion

Summary of Results

The results of the four studies are readily summarized: Extending intra-list distractor periods had little or no effect when the distractors consisted of repeated articulations of the same word. By contrast, when distractor times were extended by requiring the articulation of different words in succession, performance suffered from the added distractors. This pattern was unaffected by presentation modality of the memoranda (visual only vs. self-articulated; compare Experiment 3 and its companion study).

Figure 7 illustrates this conclusion by means of a state-trace analysis of the results from the four experiments. The figure shows, for each experiment and each serial position, the mean error rate associated with long bursts (3 or 4 distractors, depending on experiment and condition) as a function of the mean error rate associated with single-item bursts. The figure plots two functions, one for each type of burst, but averaging across other experimental variables where present (i.e., steady vs. changing in Experiment 1).

State-trace analysis relies on the idea that if a common process—such as time-based decay—underlies the effects of distractors in a complex-span paradigm, then a single monotonic function must describe the relationship between single-item bursts and longer bursts (Bamber, 1979; Loftus, Oberg, & Dillon, 2004). As is evident from the figure, the present data fail to meet that criterion: Instead, two distinct regression lines are needed to describe the relation between short and long bursts, one for the data points where the longer burst is simple and one for the points where the longer burst is complex. The two dotted lines in the figure (and associated equations) represent the best-fitting regression lines for each group of data points. The difference in their intercept reflects the observation that, across all performance levels, complex bursts lead to greater memory loss, relative to a single distractor, than simple bursts.

Statistical confirmation of this impression was provided by fitting two competing regression models to all data points: One model sought to capture the relationship between long-burst and short-burst error rates with a single intercept and slope—thus assuming that a single process captured the effects of both types of burst—and the other model differentiated between simple and complex bursts by allowing two separate intercepts (but with a common slope for both types of burst). The former model fit notably worse ($R^2 = .68$) than the latter ($R^2 = .83$) and the difference between models was highly significant, $F(1, 37) = 30.33$, $p < .0001$. (Letting slopes vary between burst types did not further improve fit; $F(1, 36) = 1.65$, $p > .1$). The state-trace analysis provides final confirmation that, across all experiments, the extent of disruption associated with complex bursts was greater than that associated with simple bursts.

There is, however, one important qualification. All experiments involved verbal distractors and memoranda and all involved forward serial recall. Our conclusions must therefore not be over-extended; for example, it is possible that a spatial distractor task might have different effects, although currently available data suggest that spatial distractor tasks affect verbal memory in qualitatively the same way as verbal distractors (Barrouillet et al., 2004, 2007; Vergauwe, Barrouillet, & Camos, 2009).

Relationship to Previous Results

Generality of results. We first place our results into a wider context by considering existing experiments in which the number of intra-list distractors was manipulated.

Figure 8 summarizes the available data from all relevant experiments known to us. The figure plots a summary statistic for the effects of manipulating time in the encoding phase through variation of the number of distractors, called the *time effect at encoding* (TEE, cf. Oberauer & Lewandowsky, 2008). The TEE represents the mean difference in performance across serial positions between long bursts (i.e., 3 or 4 distractors) and single-distractor bursts. Thus, negative values reflect a further disruption incurred by a longer delay whereas positive values represent an increase in performance.

The figure demonstrates two observations: First, across 10 experiments involving more than 200 subjects, simple bursts (white and grey bars) incur at most a small additional performance loss when their duration is extended. Second, complex bursts give rise to a very different pattern of results, causing a much greater disruption than their simple counterparts; the figure highlights this conclusion by showing data from the present experiments together with 6 additional experiments.

Dimensions of Novelty and Similarity. In a memory paradigm such as ours that combines a list of memoranda with processing of distracting material, three dimensions of similarity can be distinguished: (1) similarity between memoranda and distractors; (2) similarity among memoranda; and (3) similarity among distractors. Concerning the first dimension, most of this work comes from studies of irrelevant speech (see below) or articulatory suppression (e.g., Murray, 1967). One recent study that looked at similarity between memory items and distractors in a complex-span paradigm very similar to the present one (Oberauer, 2009). So far this line of research converges on the conclusion that similarity between memoranda and distractors has little, if any, effect on recall performance.[3] SOB has not yet been applied to the effects of similarity between memoranda and distractors.

Much research has been conducted on the second dimension, establishing that similarity between memoranda usually impairs memory for order, but under some conditions (e.g., when words rhyme) improves memory for items (e.g., Fallon, Groves, & Tehan, 1999). One of the first successes of SOB was that it explains, through novelty gating, the intricate pattern of similarity effects in lists consisting of mixtures of similar and dissimilar items (Farrell & Lewandowsky, 2003; Farrell, 2006; Lewandowsky & Farrell, 2008a).

The present studies manipulated the final dimension, namely the novelty (and hence similarity; identity is an extreme form of similarity) of each distractor relative to the other distractors on the list. We successfully tested a prediction of SOB that, again, arises from the principle of novelty-gated encoding; below, we flesh out the precise implications of our data for the theory.

The effects of irrelevant speech and articulatory suppression. Our experiments are also broadly related to work in two other areas; namely, the effects of irrelevant speech (or sound; IS from here on) and the effects of different types of articulatory suppression (AS). The effects of IS are studied by accompanying (typically visual) list presentation with an auditory stream of speech (or speech-like sounds) that the subject is instructed to ignore. The effects of AS are studied by asking subjects to articulate the irrelevant material themselves during list presentation. The similarity between the present methodology and research on the effects of IS and AS is immediately obvious: In all three cases the list is accompanied by irrelevant auditory material and emphasis is on its disruptive effects on memory performance. However, there is also at least one important difference: Whereas in the present case emphasis was on the effects of the *duration* of distracting activity in between list items, we know of no precedent in research of IS and AS that manipulated the intra-list duration of the irrelevant material. Instead, studies on the effects of IS and AS tend to present stimuli in two parallel "streams", such that articulation or speech accompanies a (typically) fairly rapidly presented list (e.g., LeCompte, 1995). This difference notwithstanding, the findings from research on IS and AS converge nicely with ours: One robust finding from work on IS is that "changing-state" sound (e.g., presentation of "A X F B D, . . . ") disrupts memory more than "steady-state" sound ("A A A . . . "). Indeed, the latter only sometimes disrupts performance (e.g., LeCompte, 1995) whereas at other times it does not (e.g., Jones, Macken, & Nicholls, 2004). The difference between steady-state and changing-state speech translates rather seamlessly into our comparison between steady-simple and changing-complex bursts, and our results thus clearly parallel the predominant effect observed with IS. Concerning AS, there is also evidence that steady-state articulation (e.g., participants repeat the syllable "be" throughout list presentation) disrupts recall less than changing-state suppression (participants repeatedly recite part of the alphabet, e.g., "A" through "G").[4]

In summary, the present studies mesh nicely with precedents in research on IS and AS. However, unlike any available precedents, our results further establish the time course of the disruption arising from the two types of distractors. Examination of that time course turns out to have significant theoretical implications.

Theoretical Implications

Revisiting time-based forgetting. What are the implications of our results on decay-based theories of short-term memory (e.g., Burgess & Hitch, 1999; Page & Norris, 1998)? Given that even simple bursts sometimes gave rise to some additional forgetting when their duration was extended (in particular in Experiment 1), might a decay view handle our data after all? We suggest that there are two grounds for dismissing this possibility. First, notwithstanding the occasional emergence of time-based forgetting with simple bursts, Figure 8 shows that overall there is little evidence to link simple bursts with time-based forgetting. Second, and more important, although a purely temporal model might handle the (limited) forgetting that we (sometimes) observed when a burst was extended from 1 to 3 (or 4) items, it is entirely unclear how the model would simultaneously handle the much more drastic forgetting that was caused by the introduction of a *single* distractor. For example, in Experiment 3, introduction of a single distractor inserted a 1.15 s interval in between items, which was accompanied by a drop in performance of nearly 30% (from .89 to .61, averaging across serial positions). Adding a further 3 repetitions of the same word extended that interval by 1.30 s, but only caused a further 2% in performance (to .59). Moreover, even if some highly non-linear decay function could be construed that fits these results, this could still not account for the difference between simple and complex distractors. We therefore suggest that decay models are not readily compatible with our results.

The problems associated with the alternative class of temporal views, namely distinctiveness models (e.g., Brown et al., 2007), are different but equally challenging. Like decay models, temporal distinctiveness models would have difficulty accommodating the simultaneous presence and absence of temporal effects for complex and simple bursts, respectively. Unlike decay models, however, the problem with distinctiveness views is that they predict an *increase* in performance with increasing intra-list distractor intervals. This is because increasing the temporal separation of items makes them more temporally distinct, and hence should render them more retrievable—a prediction that runs counter to all available data using distractors at encoding. [5]

Implications for the TBRS. We noted at the outset that TBRS can accommodate the absence of an effect of the number of distractors when cognitive load is sufficiently low to permit intermittent attentional refreshing to fully compensate the effect of decay (Panels A and B in Figure 2). However, the present studies ensured that participants processed each distractor as quickly as possible; that is, the experimenter ensured that people were articulating distractors (and to-be-remembered items in Experiments 1–3) virtually continuously, without any overt break. This manipulation pushed cognitive load to the maximum level that can be achieved with speaking as a distractor activitiy, thus strongly limiting the opportunity for intermittent refreshing of memory traces. It follows that the TBRS must predict an effect of distractor duration in the present studies (Panels C and D in Figure 2). This prediction is borne out only for complex bursts but not for simple bursts across the extant studies (see Figure 8). The absence of a distractor-duration effect for simple bursts thus constitutes a challenge to the TBRS: if there is little or no opportunity for refreshing, then the model necessarily predicts that extending the distractor duration should lead to more decay and hence further forgetting. Similarly, the large differences between simple and complex bursts are difficult to explain for the TBRS without the addition of an explicit mechanism for interference.

To escape these challenges, proponents of the TBRS might mount two counterarguments: First, it could be argued that cognitive load in our experiments was nowhere near 1, thus leaving ample time for refreshing, which would fully counteract any additional decay from additional distractors (Figure 2; Panels A and B). Second, it might be suggested that repetitions of the same word occupy the attentional bottleneck to a lesser extent than articulation of three different words, so that cognitive load is less for simple than for complex bursts. We take up those potential counterarguments in turn and evaluate them with respect to existing data.

The first counterargument, that cognitive load was comfortably below 1 so that people had sufficient time to fully compensate decay (at least in simple burst), can be sustained only if people's speaking duration grossly overestimates the time during which the attentional bottleneck is occupied. One possibility for this situation to arise is that people intentionally slowed articulation of the distractors to allow for some refreshing to take place during overt speaking (cf. Hudjetz & Oberauer, 2007). To test this possibility we conducted another experiment which was identical to Experiment 3 in nearly all respects with one important exception: Instead of memorizing and recalling the letters, participants ($N = 19$) indicated aloud whether or not a letter ended in an "ee" sound.

The sole dependent measure was the articulation time for the distractor bursts, which were measured as in Experiment 3. The results of this study are shown in Table 3 together with the corresponding articulation times from Experiment 3.

The table reveals several important points. First, articulation of distractors is generally slower (by 110 – 250 ms) when people have to commit the intervening letters to memory than when they do not. This additional time could have several sources, among them additional cognitive processes dealing with the switch between encoding memoranda and reading distractors (known as "mixing costs;" Rubin & Meiran, 2005).

Ignoring these possibilities, we could interpret the time difference as reflecting deliberate slowing of articulation in the context of a memory task to make room for refreshing. If this scenario is correct, it would imply that cognitive load in Experiment 3 was not 1 but around .9 (see the row in the table labeled CL_{min}). This calculation of cognitive load is based on the assumption, made explicit in Barrouillet, Gavens, Vergauwe, Gaillard, and Camos (2009), that the time available for refreshing is the difference between the available time for speaking a word in the complex-span experiment (i.e., the measured duration per word in our experiments) and an off-line measure of speaking duration for that word not involving memorization. A cognitive load of .9 is unlikely to leave sufficient time for refreshing to fully counteract decay to the point where it no longer makes a difference for how long the distractor burst continues. If refreshing could completely counteract decay at a load of .9, memory could hardly improve further at lower levels of cognitive load, contrary to the large effects of cognitive load that Barrouillet and colleagues have repeatedly demonstrated in the range of loads between 0.2 and 0.8. Perhaps more important, the lower-bound estimates of cognitive load (CL_{min}) computed from our auxiliary experiment do not predict memory performance. CL_{min} estimates were identical for single distractors and for complex bursts, but performance differs substantially between these conditions (see first and last rows in the table). Moreover,

CL_{min} was higher for simple bursts than complex bursts, but simple bursts resulted in better memory. Thus, if cognitive load is estimated on the basis of off-line measurements of articulation time, it mispredicts the relative order of conditions in our experiments.

We next examine the second possible counterargument, that simple bursts fail to occupy the attentional bottleneck as much as complex bursts. This is a possibility if we assume that overt speaking duration is not a good indicator of the duration of attentional capture. There is evidence consistent with the assumption that people can engage in attention-based refreshing even while overtly speaking (Hudjetz & Oberauer, 2007), so this possibility must be taken seriously. Indeed, we have entertained the possibility that simple bursts do not fully occupy attention in the past ourselves; Oberauer and Lewandowsky (2008, p. 559). At present, we have no way of measuring the duration of attentional capture by a word-reading task independent of the duration of overt speaking. If we assume that duration of attentional capture is separate from speaking duration, we can treat the former as a free parameter in the TBRS, and it is conceivable that the TBRS can account for the present results by appropriate values for this free parameter. The model would have to assume that reading a new word occupies the bottleneck for a long time to explain the large amount of forgetting by even a single distractor. At the same time, repeating the same word should capture little, if any, attention, so that in simple bursts attentional capture is much reduced after the first word of each burst. Such an account, however, will have to face two problems.

First, Oberauer and Lewandowsky (2008) showed that simple bursts during encoding do not give rise to time-based forgetting even when accompanied by a second attention-demanding symbolic choice task. People in their Experiment 2 not only had to articulate distractors in between each list item; each verbal distractor was accompanied by a speeded decision (indicate by keypress which of two characters, & or %, was flashed on the screen). Barrouillet et al. (2007) showed that choice tasks of this type occupy the attentional bottleneck. It follows that even in circumstances in which attentional capture is known to be substantial for every single distractor within a burst, extending the duration of simple bursts causes no additional forgetting. This result is incompatible with the prediction of the TBRS that the inexorable decay during attentional capture must cause time-based forgetting.

Second, Barrouillet et al. (2004) directly compared the effects of cognitive load for two types of distractors, one involving the repetition of a single syllable (i.e., "baba") and one involving the articulation of random numerals. The two types of distractors map exactly into the present distinction between simple and complex bursts, respectively. Although Barrouillet et al. (2004, Figure 2) initially reported differences between the two tasks, their chosen method of analysis was later identified as confounded (Barrouillet et al., 2007, p. 571). These authors proposed to measure processing duration and divide it by the total available processing time to obtain a more adequate estimate of cognitive load. We re-plotted the data of Barrouillet et al. (2004) as a function of cognitive load, estimated by dividing measures of speaking duration by total available time. As shown in Figure 9, "baba" span and reading-numerals span decline with increasing load by the same function. In particular the slope of span over load is equally steep for tasks with simple bursts and with complex bursts when load is measured on the basis of speaking duration. This finding is difficult to reconcile with the assumption that in simple bursts the duration of attentional capture is a smaller proportion of speaking duration than in complex bursts. If every "ba" (except the first in each burst) captured attention for a smaller proportion of overt speaking duration than reading a numeral does, then we would underestimate cognitive load for simple bursts when we compute it from speaking duration. Hence, it would be expected that the span-over-load function computed on the basis of overt speaking times should be shallower for the "baba" span than for reading-numeral span. The data of Barrouillet and colleagues in the figure show otherwise.[6]

To conclude, as long as the duration of attentional capture cannot be measured directly, the possibility remains that an account of our results can be constructed within the TBRS. We have noted two problems for such an account that we consider to be highly challenging. It is possible that they can be solved, but unless a solution is demonstrated, the balance of evidence points away from decay as the principal mechanism of forgetting in the complex-span paradigm. Our discussion of the TBRS has highlighted the importance of obtaining an independent measure of the duration of attentional capture of any task that is used as distractor task in a working-memory paradigm. Fortunately, such a procedure is available from research on the psychological refractory period (PRP; e.g., Pashler, 1994). Specifically, PRP research employs well-understood dual-task methodologies to identify if—and for how long—a task occupies the attentional bottleneck. Although it is presently unknown how readily those single-event tasks can be extended to the repeated articulation of verbal distractors, this avenue appears promising.

Finally, we must clarify that our objections to the TBRS are exclusively limited to the model's primary reliance on decay, rather than interference, to explain the interplay between forgetting and refreshing of memory traces in the complex-span task. We agree with the TBRS's premise that pauses in between distractors exert a beneficial effect upon memory, and we also agree that the distractors themselves disrupt memory. However, based on the present data and related results reported elsewhere (Lewandowsky & Oberauer, in press), we reject decay as the primary cause of that disruption and suggest instead that it is best explained by an interference mechanism. This, then, sets the stage for our concluding discussion of the SOB model introduced at the outset, which relies entirely on interference and may thus provide a preferable account of the present data.

Implications for SOB. The results of all four experiments were qualitatively in accord with the predictions of SOB (Figure 1). However, the quantitative match between theory and data varied considerably across experiments. In Experiment 3 and its companion study, the additional disruption due to complex distractors remained roughly constant across serial positions (see right-hand panels in Figure 6). In Experiments 1 and 2, in departure from SOB's predictions, that disruption diminished across serial positions and was (roughly) nil for the last item (right-hand panel in Figure 4). We noted earlier that the presence (or absence) of the final burst after the terminal item was the cause for these differences in outcome.

One might therefore wonder if SOB's predictions are also sensitive to the presence or absence of the final burst of distractors. This proved not to be the case in exploratory simulations for reasons that are readily explained: In SOB, list items and distractors alike are superimposed into a common weight matrix, using the same context marker for an item in a given serial position and all distractors in the immediately succeeding burst. At recall, when that context marker is used to probe memory, it will elicit retrieval of the item as well as the distractors in a combined ("blended") representation. Interference results because the correct response is more difficult to disambiguate from this blended representation than if no distractors were present. However, this interference is not limited to the serial position at which distractors were presented because the context markers overlap between positions: In consequence, when later (or earlier) positions are probed during recall, the blended representation that is retrieved also contains partial traces of later (or earlier) distractors, thus inheriting the interfering effect from those other positions. It follows that removal or addition of a final burst on its own has a smaller effect in SOB than in the data—indeed, the predictions in Figure 1 remain largely unchanged if a terminal burst is added to the simulation.

To resolve this problem, we introduced a modification to SOB that was independently motivated by recent empirical results. Farrell and Lelièvre (2009) compared several different representational schemes that are commonly used in models of serial recall. Of greatest interest here is their finding that the data from a number of experiments were best accommodated by a representation that combined a continuous start marker (which is roughly comparable to the context markers in SOB) with a "restricted end marker." In their modeling, the restricted end marker was implemented as a single unique signal to which the last item—and only the last item—of a list (or of a group within list) was associated. The fit of the model was improved dramatically by addition of that final marker compared to three other competing representational schemes. We therefore modified SOB to include an end marker for a final set of simulations.

Specifically, the last list item (and the subsequent distractor burst, if any), was associated to a context marker that was the weighted sum of the standard marker for the last position and a novel orthogonal marker (with weights .1 and .9, respectively). In all other respects the architecture of the model remained unchanged: Items and distractors were associated to the context markers as before, except that the similarity structure among context markers was now cognizant of the results reported by Farrell and Lelièvre (2009). The simulations inherited all parameter values from the predictions, with the exception of c, which was reduced to .35 (from .54. This change was necessary because with the original value of c, the serial position curve for the quiet baseline reached a recency-asymptote at the penultimate position, such that performance was identical for the last two items). The results are shown in Figure 10. The top panel (A) shows the simulation results without a final burst and the bottom panel (B) shows results with a final burst added. All parameter values were identical between panels; the only difference was whether or not a final burst was present.

The figure clarifies that after the addition of an independently-motivated restricted end marker, SOB's predictions mirror the data quite closely. When no final burst is present, both theory and data show near-complete convergence of all serial position curves for the last list item. When a final burst is added, the various serial position curves are widely separated for all serial positions, and the large effect of within-burst similarity extends across all serial positions. Note that with one exception (c), the parameter values in this simulation were obtained in a different setting without consideration of the present data. It would therefore be unreasonable to expect a close quantitative match between all aspects of predictions and data. In our view, the match between the predictions in Figure 10 and the data is quite striking, given the dearth of a priori quantitative predictions in the field and given the novelty of the present results.

To date, SOB has a successful and fairly long record of predicting theoretically incisive results. For example, Lewandowsky and Farrell (2008a) showed that SOB, unlike the primacy model (Page & Norris, 1998) and SIMPLE (Brown et al., 2007), could handle the advantage associated with phonologically dissimilar items on mixed lists. Likewise, Oberauer and Lewandowsky (2008) showed that SOB was best able to handle the effects of simple bursts at encoding and retrieval, and Lewandowsky, Geiger, and Oberauer (2008) showed that the model could predict the greater forgetting associated with complex bursts interspersed in between retrieval events. Because context markers evolve across events, not time, SOB is also immediately compatible with the body of research suggesting that temporal isolation effects are largely absent in forward serial recall (e.g., Lewandowsky et al., 2006; Nimmo & Lewandowsky, 2005, 2006). In this article we have extended this list of successes to include the temporal effects of various intra-list distractors. With one plausible modification based on independent evidence, SOB was able to handle the intricate interaction involving within-burst similarity, serial position, and the presence of a terminal burst of distractors.

We have also shown that purely temporal models cannot accommodate the present time-independent effects of within-burst variation (e.g., Brown et al., 2007; Burgess & Hitch, 1999). Likewise, time-based models have no mechanism to accommodate the steep decline in performance caused by a single distractor accompanied by the negligible decline after further repetitions of the same word, even if they are augmented by compensatory rehearsal (Barrouillet et al., 2004). We therefore suggest that SOB may well be developed to provide a comprehensive account of complex-span performance, including the indisputable effects of restoration time reported by Barrouillet and colleagues (e.g. Barrouillet et al., 2004, 2007).

Conclusions

We reported four experiments that provided a bridge between the predominant methodologies used in research on STM and WM. Our distractor paradigm was virtually identical to the complex-span task prominent in WM, whereas the quiet baseline conditions simultaneously instantiated the modal serial-recall paradigm in STM. We presented evidence that forgetting caused by distractors at encoding is not time-based and thus does not reflect the operation of decay. Instead, the pattern of forgetting was consistent with the predictions of an interference-based model, SOB. We conclude that WM can be best understood by building on the novelty-sensitive interference mechanism embodied in SOB and on the trade-off between forgetting and refreshing of memory traces postulated by the TBRS.

References

Anderson, J., Silverstein, J., Ritz, S., & Jones, R. (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.* New York: Oxford University Press.

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

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.

Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E., & Camos, V. (2007). Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *33*, 570-585.

Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: A time-based resource-sharing model account. *Developmental Psychology*, *45*, 477–490.

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

Brown, G. D. A., Morin, C., & Lewandowsky, S. (2006). Evidence for time-based models of free recall. *Psychonomic Bulletin & Review*, *13*, 717-723.

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

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.

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

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D., & Minkoff, S. (2002). A latent variable analysis of working memory capacity, short term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163–183.

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

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

Fallon, A., Groves, K., & Tehan, G. (1999). Phonological similarity and trace degradation in the serial recall task: When cat helps rat, but not man. *International Journal of Psychology*, *34*, 301-307.

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

Farrell, S., & Lelièvre, A. (2009). End anchoring in short-term order memory. *Journal of Memory and Language*, *60*, 209–227.

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*, 838-49.

Geiger, S. M., & Lewandowsky, S. (2008). Temporal isolation does not facilitate forward serial recall—or does it? *Memory & Cognition*, *36*, 957–967.

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.

Jones, D. M., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1008-1018.

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 656–674.

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 66–71.

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

Larsen, J. D., & Baddeley, A. D. (2003). Disruption of verbal STM by irrelevant speech, articulatory suppression, and manual tapping: Do they have a common source? *Quarterly Journal of Experimental Psychology*, *56A*, 1249–1268.

LeCompte, D. C. (1995). An irrelevant speech effect with repeated and continuous background speech. *Psychonomic Bulletin & Review*, *2*, 391–397.

Lewandowsky, S., Brown, G. D. A., & Thomas, J. L. (2009). Traveling economically through memory space: Characterizing output order in memory for serial order. *Memory & Cognition*, *37*, 181–193.

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

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

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

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

Lewandowsky, S., & Oberauer, K. (in press). No evidence for temporal decay in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Lewandowsky, S., Oberauer, K., & Brown, G. D. A. (2009). No temporal decay in verbal short-term memory. *Trends in Cognitive Sciences*, *13*, 120–126.

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

Macken, W. J., & Jones, D. M. (1995). Functional characteristics of the inner voice and the inner ear: Single or double agency. *Journal of Experimental Psychology: Leaning, Memory, and Cognition*, *21*, 436–448.

Meiser, T., & Klauer, K. C. (1999). Working memory and changing-state hypothesis. *Jouirnal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1272–1299.

Murray, D. J. (1967). The role of speech responses in short-term memory. *Canadian Journal of Psychology*, *21*, 263–276.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, *18*, 251-269.

Neath, I., & Brown, G. D. A. (2006). Further applications of a local distinctiveness model of memory. *Psychology of Learning and Motivation*, *46*, 201–243.

Neath, I., & Crowder, R. G. (1996). Distinctiveness and very short-term serial position effects. *Memory*, *4*, 225-242.

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

Nimmo, L. M., & Lewandowsky, S. (2006). Distinctiveness revisited: Unpredictable temporal isolation does not benefit short-term serial recall of heard or seen events. *Memory & Cognition*, *34*, 1368–1375.

Oberauer, K. (2009). Interference between storage and processing in working memory: Feature overwriting, not similarity-based competition. *Memory & Cognition*, *37*, 346–357.

Oberauer, K., & Lange, E. B. (2008). Interference in verbal working memory: Distinguishing similarity-based confusion, feature overwriting, and feature migration. *Journal of Memory and Language*, *58*, 730–745.

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review*, *115*, 544–576.

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*(2), 220-244.

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

Portrat, S., Barrouillet, P., & Camos, V. (2008). Time-related decay or interference-based forgetting in working memory? *Journal of experimental psychology: Learning, memory, and cognition*, *34*, 1561–1564.

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task-switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1477–1491.

Saito, S., & Miyake, A. (2004). On the nature of forgetting and the processing-storage relationship in the reading span performance. *Journal of Memory and Language*, *50*, 425-443.

Toppino, T. C., & Pisegna, A. (2005). Articulatory suppression and the irrelevant speech effect in short-term memory: Does the locus of suppression matter? *Psychonomic Bulletin & Review*, *12*, 374–379.

Towse, J. N., Hitch, G. J., & Hutton, U. (2000). On the interpretation of working memory span in adults. *Memory & Cognition*, *28*, 341-348.

Turner, M., & Engle, R. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *49*, 446–468.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilitites. *Psychological Bulletin*, *133*, 1038-1066.

Vergauwe, E., Barrouillet, P., & Camos, V. (2009). Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1012–1028.

Wilson, M. D. (1988). The MRC Psycholinguistic Database; Machine readable dictionary, version 2. *Behavioral Research Methods, Instruments, and Computers*, *10*, 6-11.

Appendix

Details of SOB Simulations

SOB has been presented previously in some detail (e.g., Farrell, 2006; Lewandowsky & Farrell, 2008a; Lewandowsky, Geiger, & Oberauer, 2008; Oberauer & Lewandowsky, 2008). The complete source code (in MatLab) of the present simulations is available at the web page provided in the Author Note. A full exposition of the theory in a methodologically similar context can be found in the Appendix of Lewandowsky, Geiger, and Oberauer (2008). Here, we only detail the item and distractor representations which are unique to this study.

All study items in the present simulation were derived from a common item prototype (a binary vector that was randomly sampled for each simulation replication) by retaining each feature with probability p_c and re-sampling it (from a symmetric binary distribution) with probability $1 - p_c$. The value of p_c was fixed at .7, which captures the mean similarity among vectors that instantiate a multi-dimensional scaling solution of the actual stimuli (for details see Lewandowsky & Farrell, 2008a).

Distractors, when present, were associated with the positional marker used for encoding of the immediately preceding list item and were encoded in the same novelty-sensitive way as list items. 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 re-sampling it (from a symmetric binary distribution) with the complementary probability. To reflect the obvious fact that distractors differed more from the ensemble of list items than they differed amongst 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 for each burst by retaining features of the prototype with probability J_c and re-sampling 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, each distractor 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.) Between-burst variation 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 (2008). 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 list positions, the internal representation of those words likely changed somewhat across serial positions.

The predictions of SOB in Figure 1 were thus based on three fixed parameters; $t_c = 0.5$, $p_c = 0.7$, and $s_c = 0.35$, and 5 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 (2008, their Experiment 1). Specifically, $\phi_e = 300$, $\phi_s = 0.61$, $N_o = 1.38$, $c = 0.54$, and $J_c = 0.31$. The predictions were based on 5,000 replications per condition. The only change made to the simulation from the published precedent (Oberauer & Lewandowsky, 2008) was a reduction in list length from 6 to 5 and the elimination of the terminal burst (i.e., distractors following the last list item) in order to align the predictions exactly with the methodology of the present Experiment 1. For parsimony, the simulations also omitted the "refreshing" (i.e., re-encoding of the present list item after articulating each distractor) that was used in some of the simulations of Oberauer and Lewandowsky (2008); the predictions are not materially altered if refreshing is turned on.

Author Note

Preparation of this paper was facilitated by a Discovery Grant from the Australian Research Council to the first and fourth author, and an Australian Professorial Fellowship to the first author. Klaus Oberauer was also supported by a grant from the UK Economic and Social Research Council (ESRC), grant RES-062-23-1199. Address correspondence to the first author at the School of Psychology, University of Western Australia, Crawley, W.A. 6009, Australia. Electronic mail may be sent to lewan@psy.uwa.edu.au. Personal web page: http://www.cogsciwa.com. We thank Charles Hanich and Abel Oh for assistance during data collection, and Simon Farrell for comments on an earlier draft of this article. We also thank Pierre Barrouillet for a detailed critique of our conclusions.

Footnotes

[1] The term "interference" is often used descriptively (i.e., to characterize the observed effects of distractors) as well as theoretically (i.e., to refer to an underlying cognitive process), thus creating some potential for ambiguity. We seek to avoid that ambiguity by using "disruption" as the descriptive term while reserving the term "interference" for theoretical considerations only.

[2] Readers may note the similarity between this procedure and research on the irrelevant-speech effect and the effects of different types of articulatory suppression. Those empirical and theoretical links are explored in the General Discussion, once all the data have been presented.

[3] The study by Oberauer (2009) suggests that a distinction must be made between similarity and degree of feature overlap (see also Oberauer & Lange, 2008) but this distinction is not relevant here.

[4] The greater disruption with changing-state AS has been reported repeatedly (Jones, Farrand, Stuart, & Morris, 1995; Macken & Jones, 1995; Meiser & Klauer, 1999; Toppino & Pisegna, 2005). To our knowledge, only one study (Larsen & Baddeley, 2003) has failed to replicate the difference between steady-state and changing-state AS. However, the results of Larsen and Baddeley must be taken with some caution because they relied on a between-experiment comparison. On balance, the available evidence involving AS thus also parallels the distinction between steady-state and changing-state speech.

[5] One possible solution to this problem involves the idea of linking the discriminability of items (represented by a parameter within the SIMPLE model of Brown et al., 2007) to the overall time scale of list presentation. Oberauer and Lewandowsky (2008) explored this solution using a formalism suggested by Neath and Brown (2006), but found that while it can eliminate the prediction that longer intervals should lead to better memory, it cannot reverse the prediction. That is, SIMPLE can at best predict

that distractors at encoding should have no effect at all compared to the quiet baseline; it cannot, however, predict the detrimental effects observed in the present study (in particular with complex distractors). We therefore conclude that distinctiveness models are also at odds with our data.

[6] Two further questions concerning Figure 9 are worth considering: (a) How can we be certain that declining cognitive-load functions are suitable for affirming that a task engages the attentional bottleneck? In response, Barrouillet et al. (2007) showed that *flat* cognitive-load functions result for tasks that are known—on the basis of an independent assay—not to engage the bottleneck. Only tasks that are attentionally demanding occupy the bottleneck and produce a cognitive-load function with a steep negative slope. It follows that it is difficult to deny that the "baba" span task captured attention, which entails the strong implication that our simple bursts likewise occupied the attentional bottleneck. (b) How can the pattern in the figure be reconciled with the outcome of our experiments? If simple and complex bursts differ so dramatically in our studies, why do they fall on the same cognitive-load function? The answer lies in the fact that the maximum cognitive load for the "baba" span in the figure was .5—a relatively low level of load for which the TBRS need not predict a duration effect because restoration may fully compensate for temporal decay (see Panels A and B in Figure 2). This opportunity for restoration may equally counteract the effects of interference; hence SOB's predictions, which were obtained when cognitive load was assumed to be 1, may also differ for lower loads although this remains to be ascertained once a restoration mechanism for SOB has been developed.

Table 1

Mean Articulation Times (and Standard Deviations) for Experiment 1.

		Between-burst variation			
Within-burst variation	N[a]	Steady		Changing	
Simple	1	.96	(.30)	1.09	(.36)
	3	1.78	(.33)	1.91	(.40)
Complex	1	.90	(.21)	1.07	(.32)
	3	1.89	(.38)	2.26	(.60)

[a] N refers to the number of distractors in a burst.

Table 2

Summary of the omnibus ANOVA on recall performance for Experiment 1.

Effect	F^a	MSE	p^b	η_p^2
Within-burst variation (WB)	.16	.42		.00
Between-burst variation (BB)	79.98	.03	$< .0001$	.68
WB × BB	.90	.03		.03
Number of distractors (ND)	35.47	.03	$< .0001$	.48
WB × ND	4.59	.03	$< .05$	.10
BB × ND	4.82	.02	$< .05$	.11
WB × BB × ND	.22	.02		.01
Serial position (SP)	45.56	.05	$< .0001$	.55
WB × SP	1.72	.05		.04
BB × SP	2.80	.01	$< .05$	.07
WB × BB × SP	1.39	.01		.03
ND × SP	5.88	.02	$< .0002$	.13
WB × ND × SP	1.34	.02		.03
BB × ND × SP	.93	.02		.02
WB × BB × ND × SP	1.01	.02		.02

[a] *df* are 1, 38 for all effects except those involving SP, for which df=4, 152.

[b] All $p > .1$ are omitted.

Table 3

Mean Articulation Times (and Standard Errors) in ms *for Distractor Bursts With and Without Memory Demand.*

	Type of burst			
Source	Quiet	1 distractor	4 simple	3 complex
With memory demand (Experiment 3)		1152 (27.7)	2452 (58.1)	2418 (72.9)
Without memory demand [a]		1042 (16.2)	2282 (49.7)	2165 (64.8)
Difference		110	170	253
CL_{min} [b]		.90	.93	.90
Mean correct recall (Experiment 3)	.89	.61	.59	.45

[a] Data from control experiment ($N = 19$) in which letters were not memorized but subjected to a rhyming judgment.

[b] CL_{min} refers to the lower-bound estimate of cognitive load in Experiment 3, computed from articulation times with and without memory demand.

Figure Captions

Figure 1. Correct-in-position serial position curves predicted by SOB for the distractor manipulations explored in this article (Q refers to a distractor-free quiet baseline condition, 1E and 3E refer to 1 and 3 distractors, respectively, in between each pair of study items; the same labeling is used in all figures). The upper panels show predictions for simple bursts and the lower panels for 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.

Figure 2. Time-line of events and predictions of TBRS during one processing episode in between two memory items (M1 and M2) in a complex-span trial. Panels A and B show the effects of adding distractors when cognitive load is low (e.g., Barrouillet et al., 2004) and Panels C and D show the effects of the same manipulation when cognitive load is approaching 1 (as in the present studies). In all panels, distractors are represented by downward pointing triangles (labeled P) and are accompanied by decay. Restoration times, which are available for attentional refreshing when cognitive load is low, are shown by upward pointing triangles (R). In each panel, the predicted evolution of memory strength of item M1 is shown below the events. In Panels A and B, each distractor is followed by compensatory attentional refreshing, thus fully reversing the effects of decay. As a result, adding further distractors (Panel B) does not cause a decline in performance. In Panels C and D, cognitive load is 1 and hence no time is available for refreshing. Additional distractors (Panel D) therefore cause additional decay and must lead to a performance decrement. Predictions in Panels C and D hold also when some minimal restoration time is available and cognitive load is minimally below 1. See text for further details.

Figure 3. Serial Position curves obtained for all conditions in Experiment 1. The proportion correct-in-position recall is shown separately for the 4 different conditions, with rows showing the effect of within-burst variation and columns the effect of between-burst variation. The upper panels show data from the simple-burst group, the lower panels show data from the complex-burst group. The left-hand panels show the effects of steady bursts and the right-hand panels the effects of changing bursts. Note that the quiet baseline data (Q) are identical for both panels within each row because that condition was not fully crossed with the between-burst variation manipulation, unlike the lists with one (1E) and three (3E) distractors. Error bars represent within-subject standard errors; see text for more explanation.

Figure 4. The effects of within-burst variation in Experiment 1. The left-hand panels show the effects 1 or 3 distractors (1E and 3E, respectively) in comparison to baseline (Q) for simple bursts (top) and complex bursts (bottom) averaging across between-burst variation. The right-hand panel directly compares the effects of adding further distractors to the two types of bursts by plotting the performance difference between 1E and 3E for the two types of bursts.

Figure 5. Correct-in-position serial position curves for Experiment 2. Error bars represent within-subject standard errors.

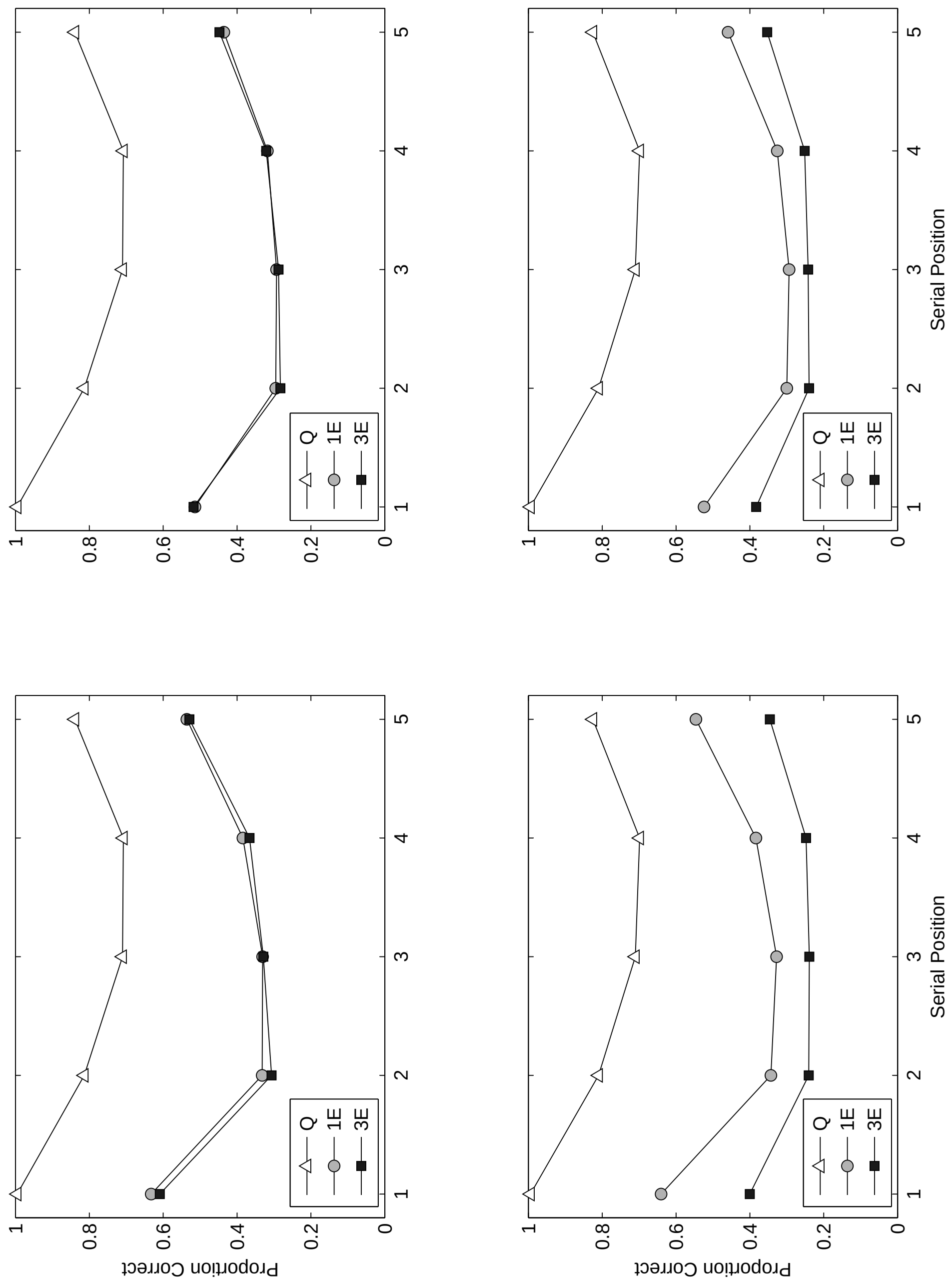
Figure 6. The left panel shows correct-in-position serial position curves for Experiment 3. Error bars represent within-subject standard errors. The right panel shows the differences between 1 and 3 distractors (for complex bursts), or between 1 and 4 distractors (for simple bursts) across serial positions. The large circles on the right represent the mean across serial positions.

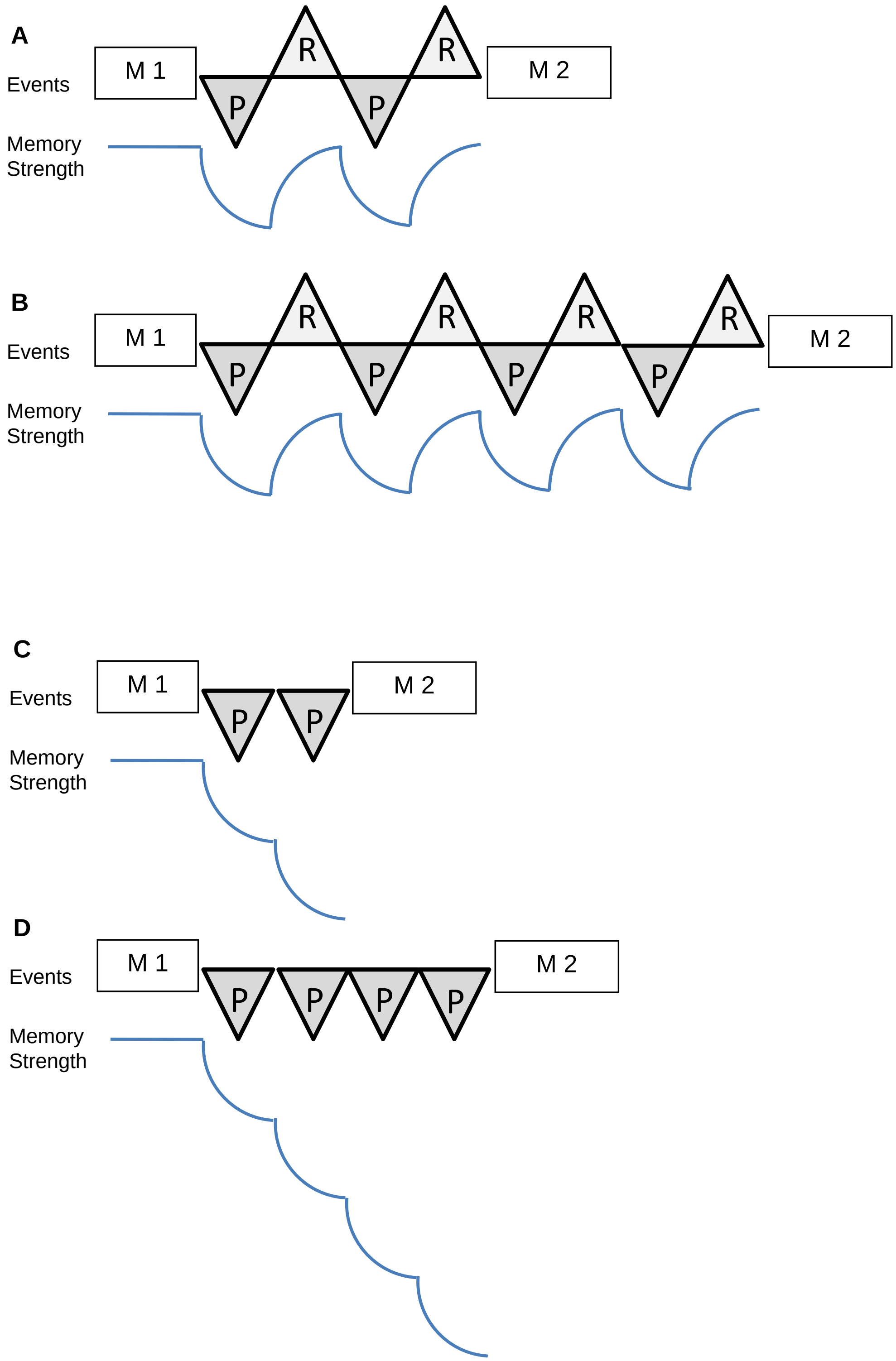
Figure 7. State-traces of simple and complex bursts. Error rates with long bursts (3 or 4 words spoken aloud) are plotted as a function of error rates after single-item bursts for all serial positions in all experiments. The two lines represent the two best-fitting slope estimates for simple and complex bursts. Size of plotting symbols is proportional to observed encoding times for long bursts (i.e., 3 or 4 words). See text for details.

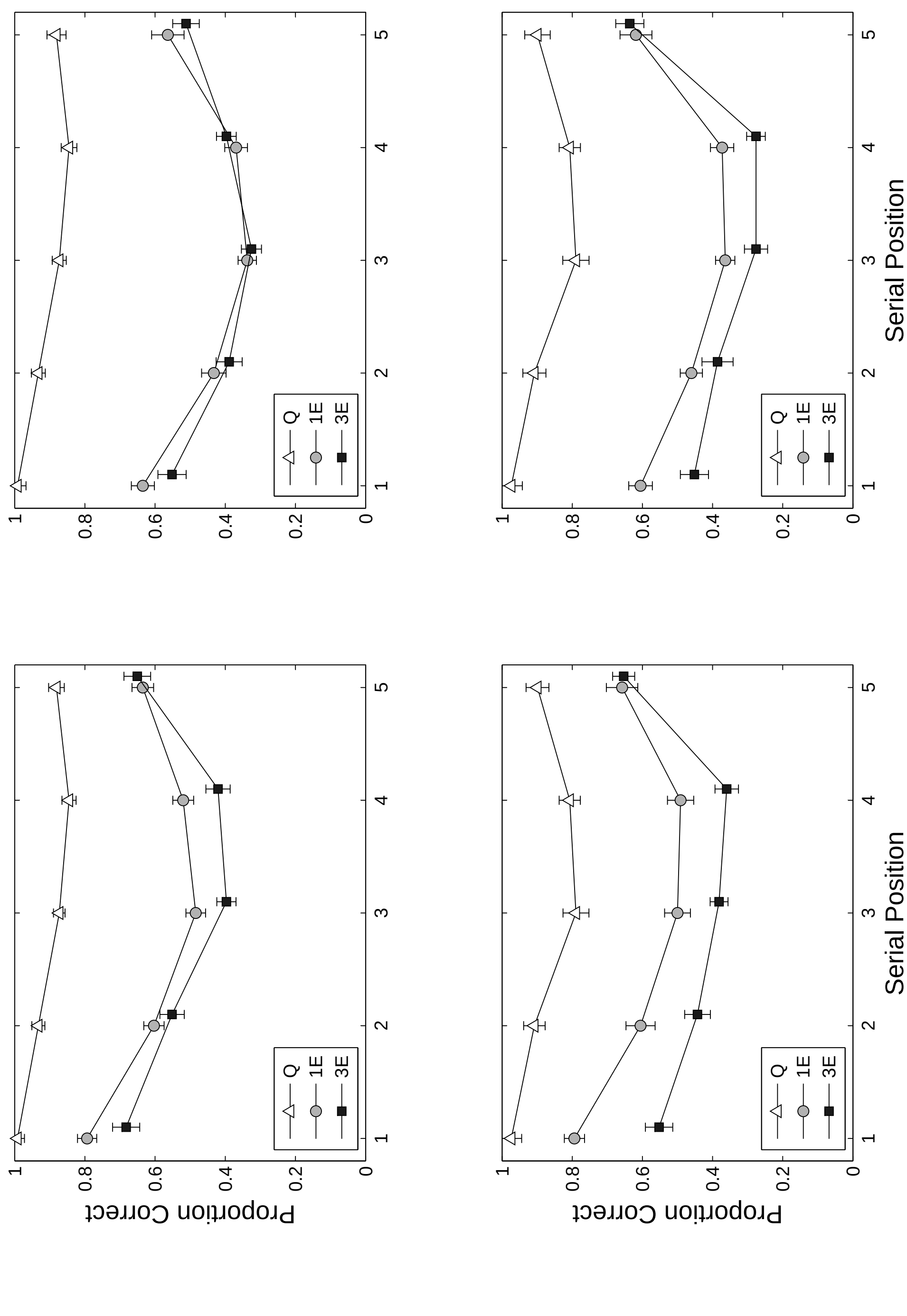
Figure 8. Summary of all existing time effects at encoding (TEE). Each bar represents a separate experiment or condition. The black bars represent the results from the present experiments with complex bursts. The remaining bars involve simple bursts; the gray bars are from the present experiments and the white bars from Oberauer & Lewandowsky (2008)'s meta-analysis (see their Table 4). The dashed vertical line represents the mean TEE across all experiments with simple bursts, and the dotted vertical lines represent the associated 95% confidence interval. See text for further details.

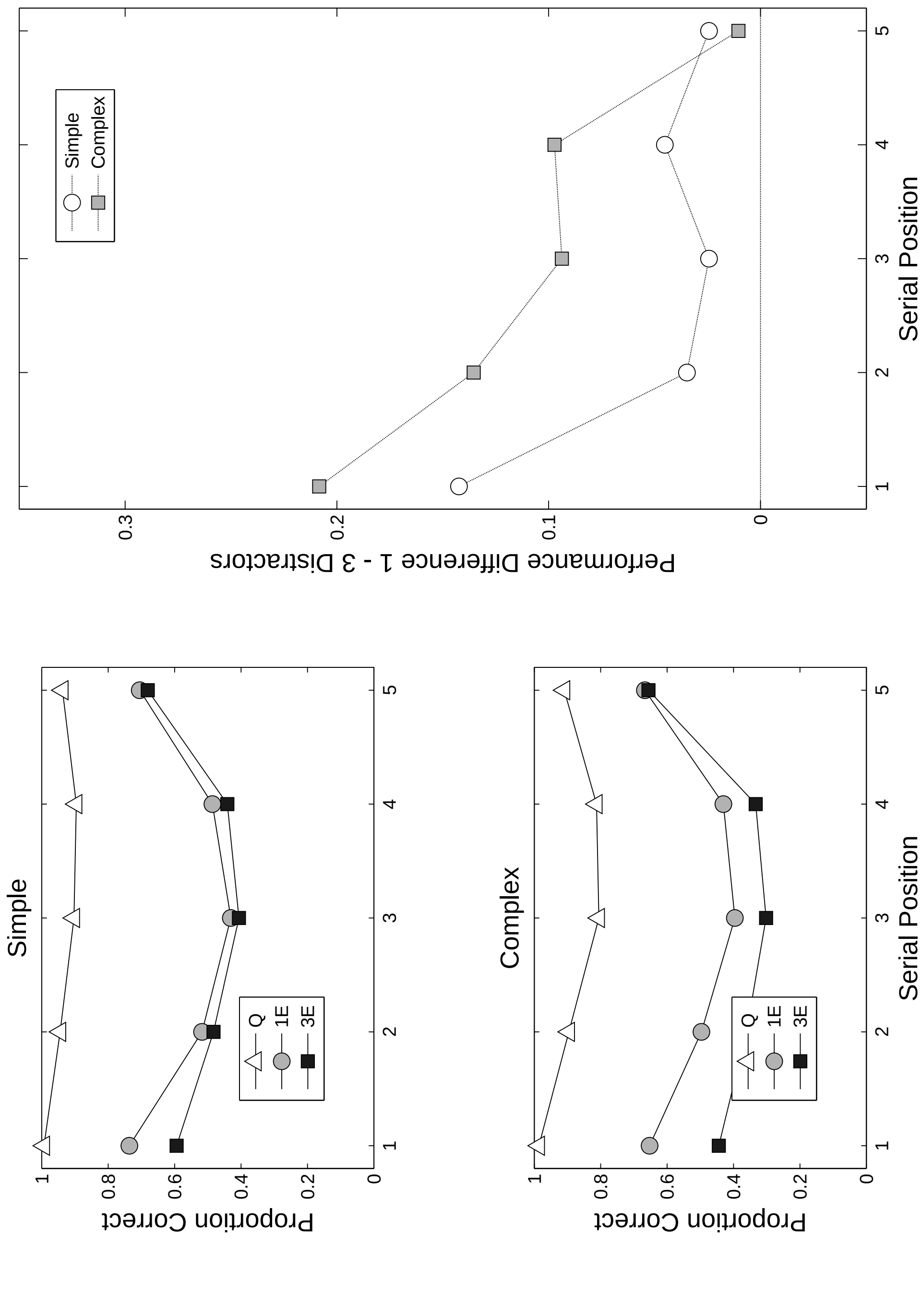
Figure 9. Cognitive load functions first reported by Barrouillet et al. (2004) replotted using an unconfounded measure of cognitive load. As suggested by Barrouillet et al. (2007), cognitive load is best expressed as the ratio of processing time and the total time available. The total times were determined by the methodology of Barrouillet et al. (2004, Experiment 7), and processing times were determined empirically. Reading of numerals has been measured to require 424 ms (Barrouillet et al., 2007, Experiment 1), and articulation of "ba" was measured to require 260 ms (data from our laboratory). See text for further details.

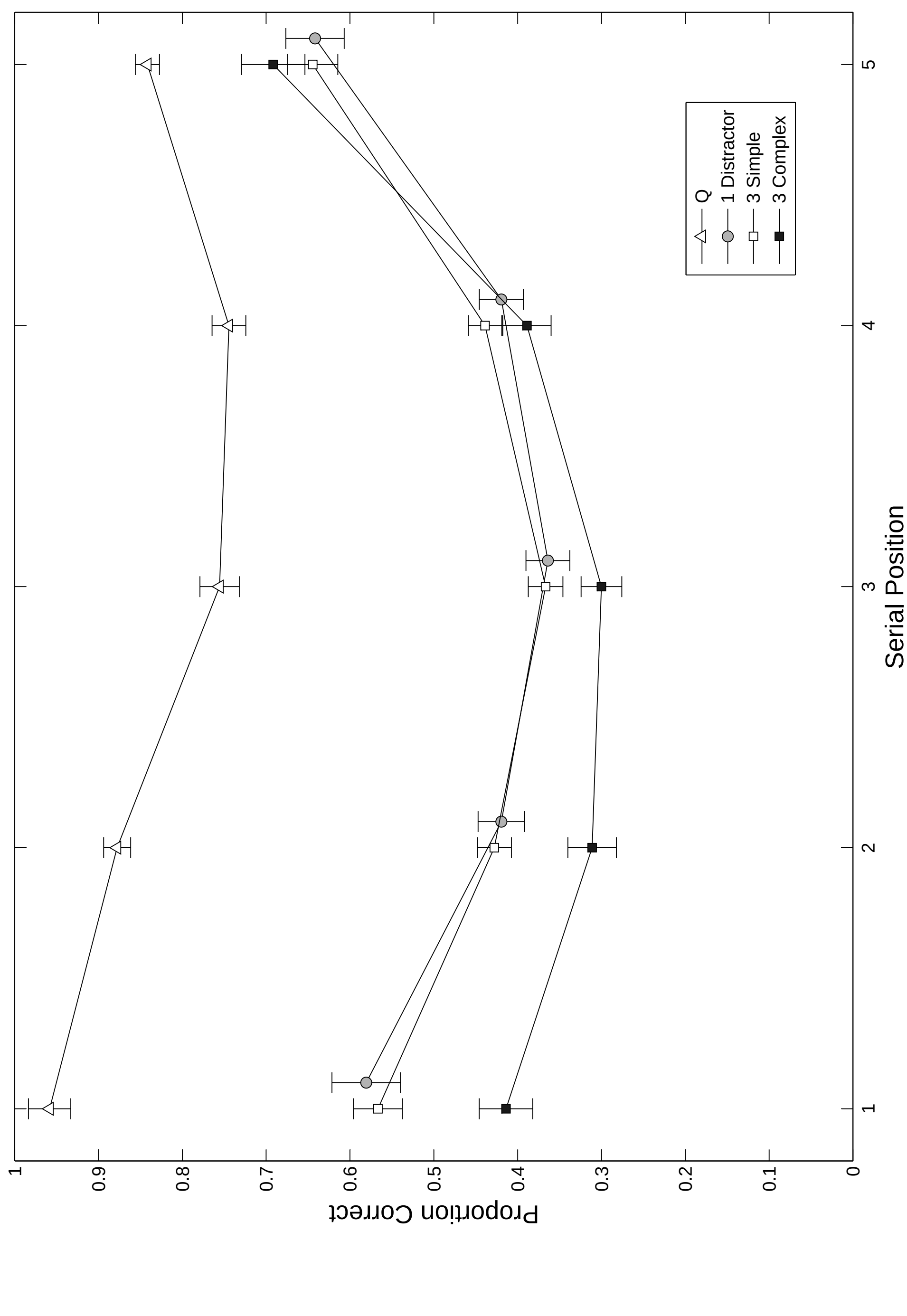
Figure 10. Simulations of SOB with a distinct end marker. All parameters and assumptions are identical across panels. Panel A shows the results without a final burst (as in Experiments 1 and 2) and Panel B shows the results with a final burst (as in Experiment 3). See text for details.

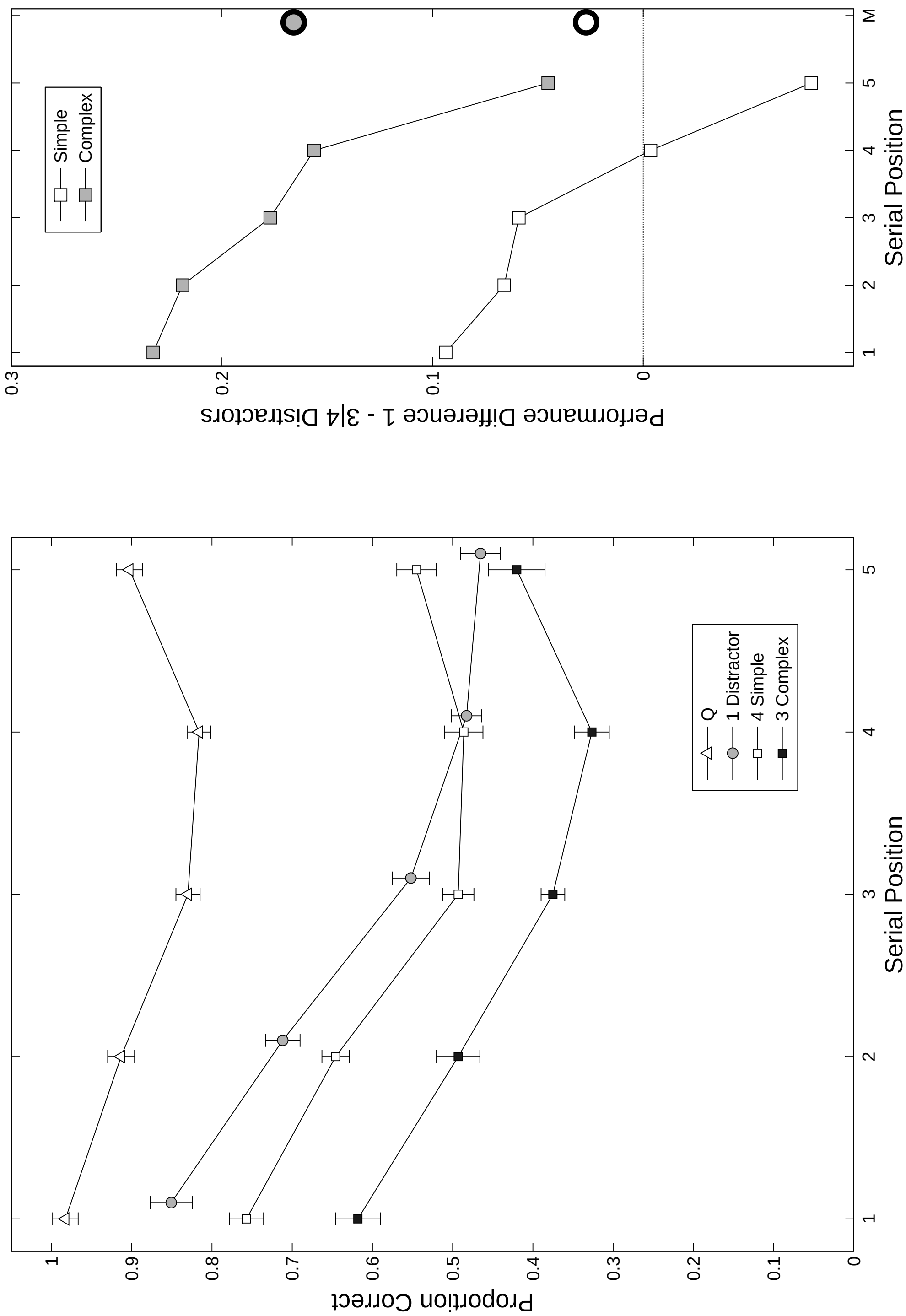

Proportion Correct
Serial Position
Q
1E
3E
0
0.2
0.4
0.6
0.8
1
1
2
3
4
5

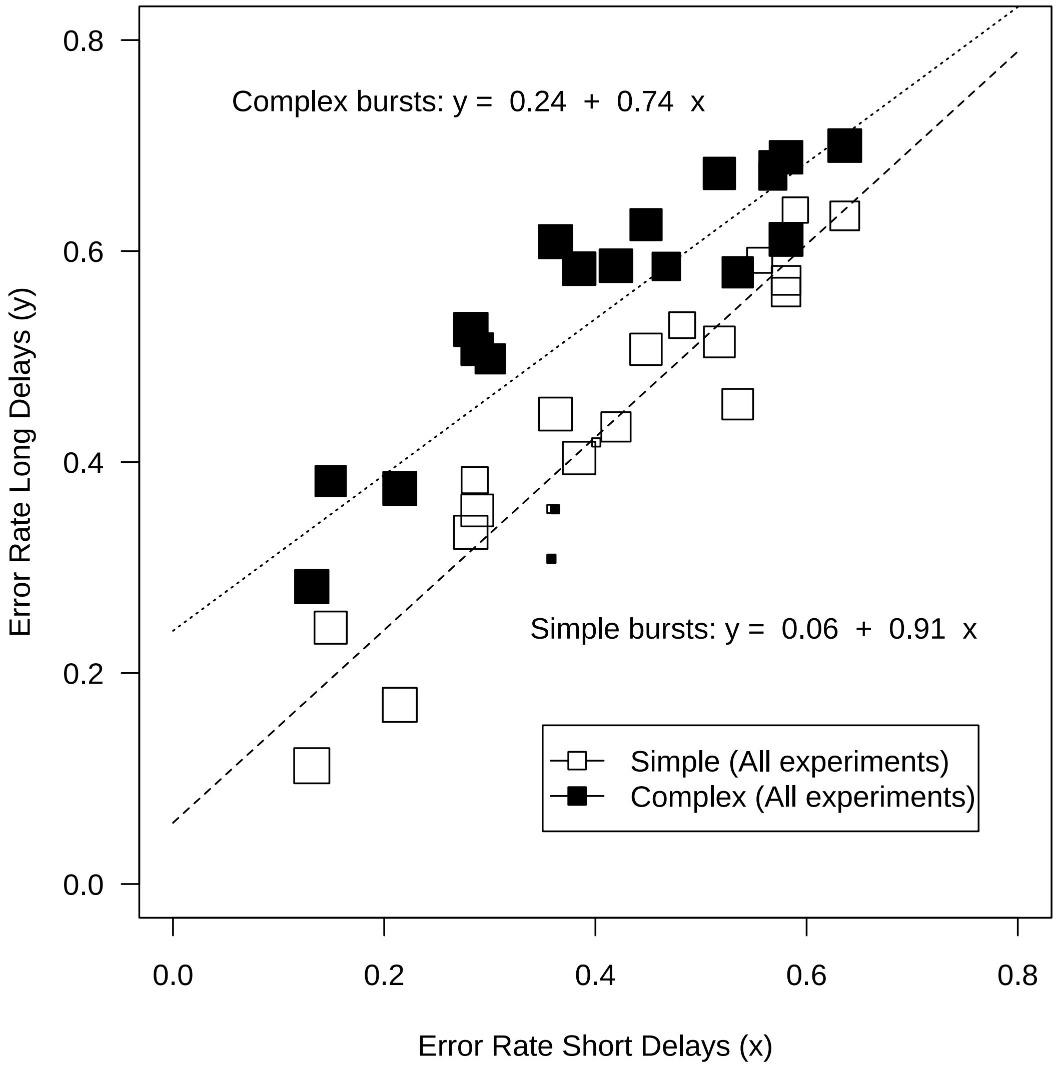

A
Events
M 1
R
R
P
P
M 2
Memory Strength
B
Events
M 1
R
R
R
R
P
P
P
P
M 2
Memory Strength
C
Events
M 1
P
P
M 2
Memory Strength
D
Events
M 1
P
P
P
P
M 2
Memory Strength

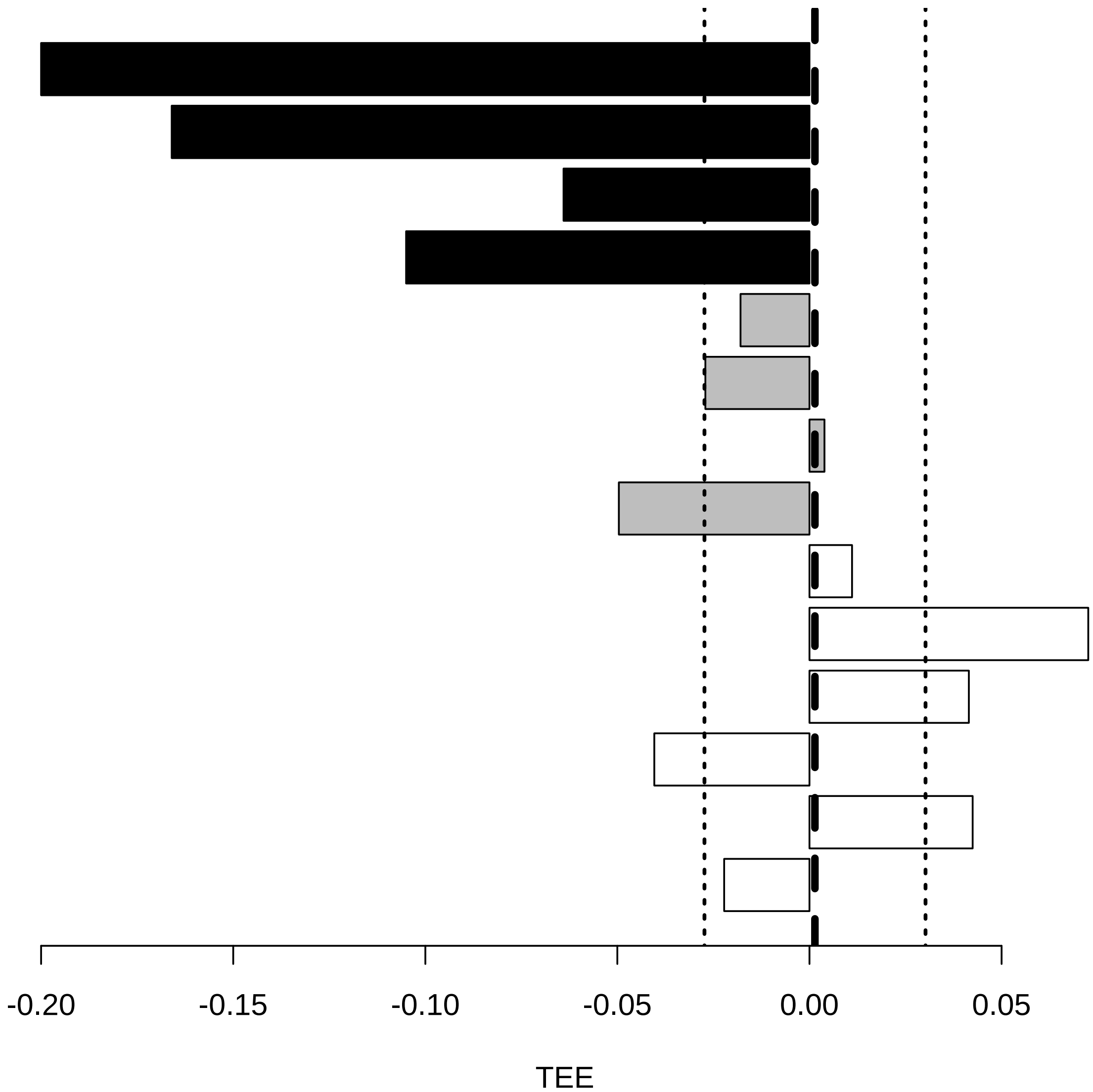

Proportion Correct
Serial Position
Q
1E
3E
Proportion Correct
Serial Position
Q
1E
3E
Q
1E
3E
Q
1E
3E

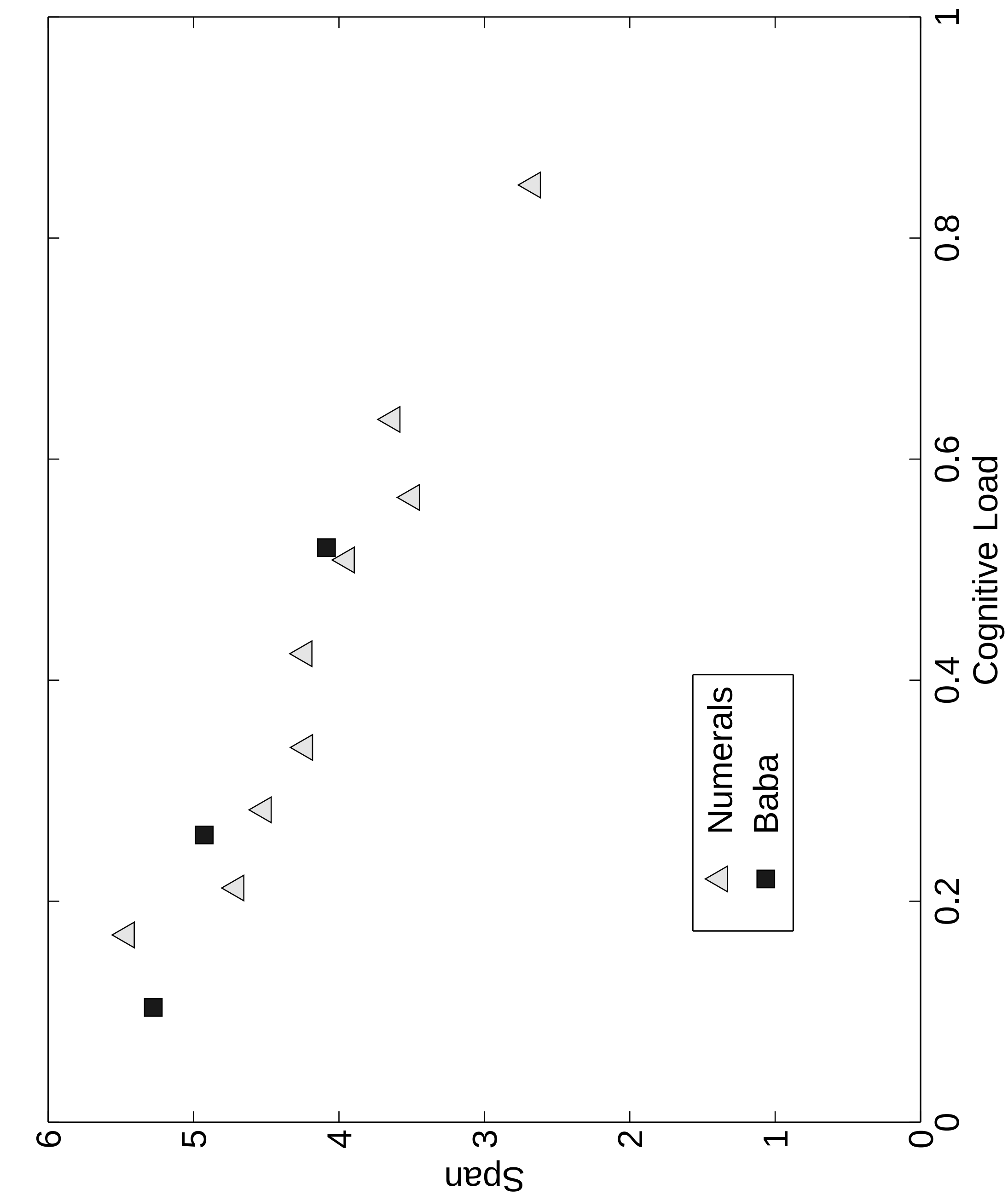

Simple
Complex
Proportion Correct
Serial Position
Q
1E
3E
Performance Difference 1 - 3 Distractors
Simple
Complex

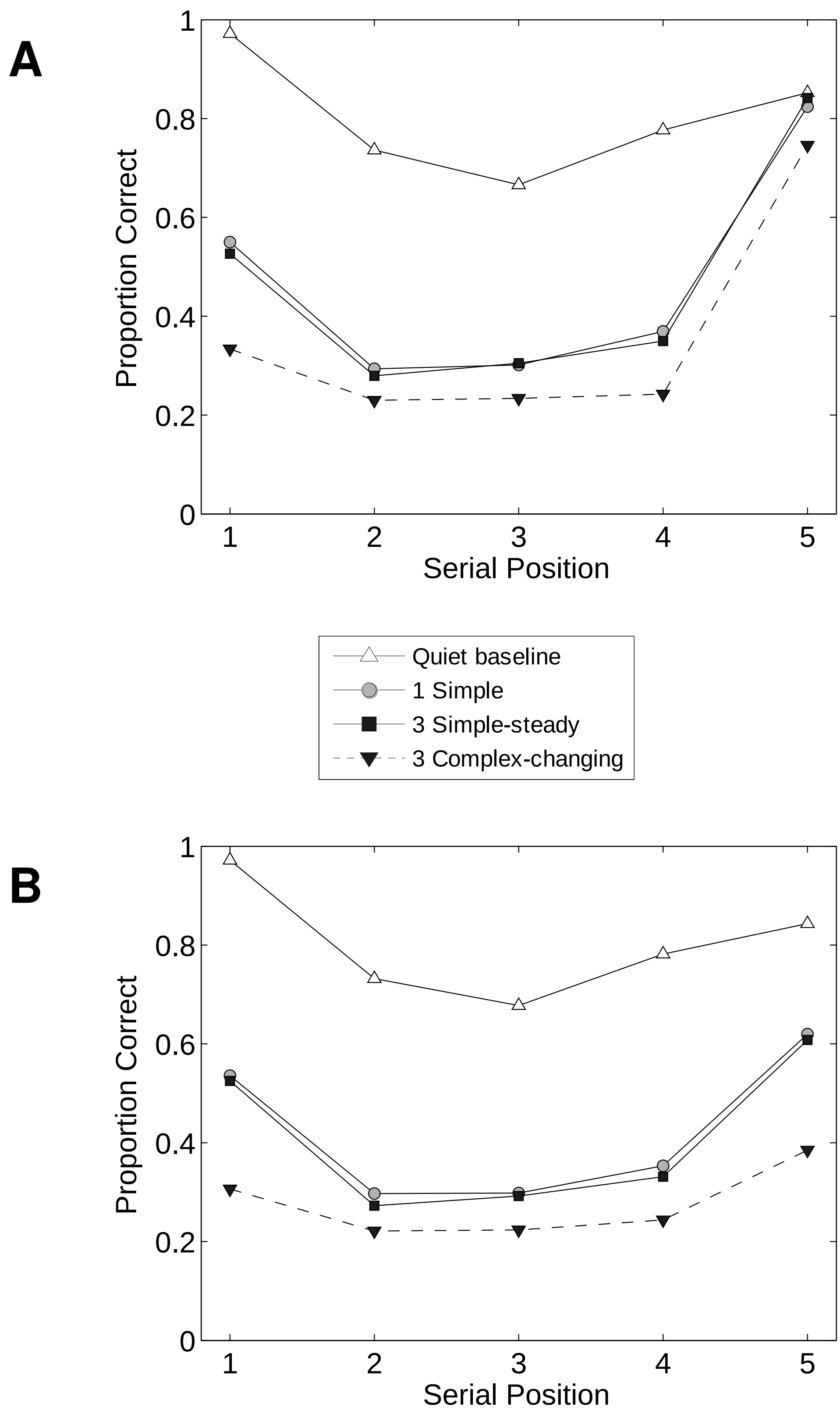

Proportion Correct
Serial Position
Q
1 Distractor
3 Simple
3 Complex
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1
2
3
4
5


Proportion Correct
Serial Position
Q
1 Distractor
4 Simple
3 Complex
Performance Difference 1 - 3|4 Distractors
Serial Position
Simple
Complex


Complex bursts: y = 0.24 + 0.74 x
Simple bursts: y = 0.06 + 0.91 x
Simple (All experiments)
Complex (All experiments)
Error Rate Long Delays (y)
Error Rate Short Delays (x)
0.0
0.2
0.4
0.6
0.8


-0.20
-0.15
-0.10
-0.05
0.00
0.05
TEE


Span
Cognitive Load
Numerals
Baba
0
1
2
3
4
5
6
0.2
0.4
0.6
0.8


A
Proportion Correct
Serial Position
Quiet baseline
1 Simple
3 Simple-steady
3 Complex-changing
B
Proportion Correct
Serial Position