

THEORETICAL AND REVIEW ARTICLES

Time does not cause forgetting in short-term serial recall

STEPHAN LEWANDOWSKY and MATTHEW DUNCAN

University of Western Australia, Crawley, Western Australia, Australia

and

GORDON D. A. BROWN

University of Warwick, Coventry, England

Time-based theories expect memory performance to decline as the delay between study and recall of an item increases. The assumption of time-based forgetting, central to many models of serial recall, underpins their key behaviors. Here we compare the predictions of time-based and event-based models by simulation and test them in two experiments using a novel manipulation of the delay between study and retrieval. Participants were trained, via corrective feedback, to recall at different speeds, thus varying total recall time from 6 to 10 sec. In the first experiment, participants used the keyboard to enter their responses but had to repeat a word (called the *suppressor*) aloud during recall to prevent rehearsal. In the second experiment, articulation was again required, but recall was verbal and was paced by the number of repetitions of the suppressor in between retrieval of items. In both experiments, serial position curves for all retrieval speeds overlapped, and output time had little or no effect. Comparative evaluation of a time-based and an event-based model confirmed that these results present a particular challenge to time-based approaches. We conclude that output interference, rather than output time, is critical in serial recall.

The question of whether forgetting occurs because of time or because of interference is one of the longest-standing issues in memory research (e.g., McGeoch, 1932; Waugh & Norman, 1965). All other things being equal, does it matter how much time has elapsed between encoding and retrieval of an event? What are the effects of delaying retrieval? The role of time in memory is still hotly debated (e.g., Brown & Chater, 2001; Nairne, 2002), with proponents arguing that time and memory are closely and crucially linked (Brown & Chater, 2001) and others submitting evidence against this proposition (e.g., Friedman, 2001; Henson, 1999; Nairne, 2002). Here we address the issue of time-based versus interference-based forgetting in the context of short-term serial recall. Some models of short-term serial recall ascribe a central role to the passage of time in the recall process. The effects of time during output often underlie key properties of short-term serial

recall models, such as their ability to account for word length effects, output order effects, and the extended primacy in the serial position curve. However, many findings that have been cited in support of the role of time can also be explained by other factors, such as interference during recall or the particular nature of materials used, thus leaving the issue unresolved.

This article further explores the role of time in forgetting in two experiments that manipulated the timing of retrieval while keeping all other variables constant. Two contrasting predictions were derived by simulation: According to *time-based models*, the passage of time during serial recall is crucial to information loss, whereas *event-based models* state that forgetting occurs because of some other process such as interference from recall events. Both experiments found evidence that was consistent with event-based models but inconsistent with time-based forgetting.

TIME AND FORGETTING: THEORY

Time-Based Models

In the class of time-based memory models (e.g., Baddeley, 1976, 1986; Bjork & Whitten, 1974; Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Crowder, 1976; Glenberg & Swanson, 1986; Neath, 1993; Page & Norris, 1998), the delay between encoding and retrieval is a crucial

This research was supported by a Large Grant from the Australian Research Council to S.L. and an associated Linkage Grant awarded to S.L. and G.D.A.B., and by Grant 88/S15050 from the Biotechnology and Biological Sciences Research Council (U.K.) and Grants R000239002 and R000239351 from the Economic and Social Research Council (U.K.). We thank Leo Roberts for his assistance during data collection and manuscript preparation. Correspondence concerning this article should be addressed to S. Lewandowsky, School of Psychology, University of Western Australia, Crawley, WA 6009, Australia (e-mail: lewan@psy.uwa.edu.au).

determinant of memory performance. All other things being equal, time-based models necessarily predict that an event will become less retrievable, the longer the time that has elapsed since its encoding.

The underlying mechanisms of memory models are heterogeneous. Decay models (e.g., Baddeley, 1986; Page & Norris, 1998) postulate that encoded information “fades away” within a few seconds unless (overt or subvocal) rehearsal is used to “refresh” memory traces. An immediate implication of decay is that output time necessarily affects memory, in such a way that the longer it takes to retrieve an item under otherwise identical conditions, the poorer the recall will be. Indeed, at first glance, there is much evidence to support that contention (e.g., Cowan et al., 1992; Cowan, Nugent, Elliott, & Geer, 2000). However, Nairne (2002) recently compiled strong evidence against the decay view; among the many contrary findings that he reviewed, perhaps the most crucial is that memory performance sometimes *increases* with increasing delay (e.g., Turvey, Brick, & Osborn, 1970). For the purposes of this article, we therefore do not rely on trace decay as a guiding theoretical principle. Critically, this decision does not imply the simultaneous dismissal of chronological time as an explanatory principle. A number of theories reject decay but nonetheless nominate time as a core principle of memory, one that is responsible for forgetting. These *temporal distinctiveness* models are central to our discussion of time-based approaches, because they provide a straightforward but explicit framework within which quantitative predictions may be derived.

One computational instantiation of temporal distinctiveness, known as SIMPLE (scale invariant memory, perception, and learning), was recently provided by Brown, Neath, and Chater (2002; see also Brown & Chater, 2001; Lewandowsky & Brown, in press). We use the model to illustrate the effects of time-based forgetting during forward serial recall. In SIMPLE, items are confusable with other items in memory, and hence difficult to retrieve, to the extent that they are temporally close when considered from the retrieval perspective. More specifically, the confusability between any two memory items, $S_{i,j}$, is related to the (exponentiated) ratio of the elapsed time from their encoding to the time of recall. This mechanism favors recent items over more distant events. For example, events that have occurred 1 and 2 sec ago are less confusable ($.5^c$) than those from 5 and 6 sec ago ($.83^c$), where c is the main free parameter of the model. Because recall probability is inversely related to confusability, items from further in the past will be more difficult to recall (see Brown et al., 2002, for further details).¹

Like decay models, the purely temporal version of SIMPLE predicts that recall accuracy should decline as the time between encoding and retrieval of an item is increased. This characteristic can be clarified by comparing two hypothetical conditions in which the time between successive recalls is either short or long. Performance on the first item should be identical for both conditions be-

cause time since encoding would be equivalent (assuming identical retention intervals). However, as recall proceeds, performance on each successive list item in the long condition should get worse for the reasons cited earlier: Specifically, consider the distinctiveness ratio for, say, the last two of six list items (presented 1 sec apart) at the point when the fifth item is to be recalled. Assuming a recall speed of 1 sec/item and 3 sec/item for the short and long conditions, respectively, the corresponding distinctiveness ratios (with a 0-sec retention interval) would be $4/5^c$ ($= .80^c$) and $12/13^c$ ($= .92^c$). Hence, performance in the long condition should gradually diverge from performance in the short condition, and the serial position curves between two such conditions should show fanning out from a common origin. Time-based forgetting, therefore, expresses itself not as a main effect of recall speed, but as a fanning out that is captured by an interaction of speed with serial position. The importance of fanning cannot be overemphasized, and we illustrate this prediction with detailed simulations below.

The prediction that different recall speeds should lead to fanning of serial position curves is not unique to SIMPLE: Any model that acknowledges a similar role of time in memory would expect recall to be better for items that are recalled in quick succession than for others whose retrieval is separated by longer pauses, and that effect would also increase with the number of items that are recalled. Such models are too varied to review in detail here, but whether time-based forgetting results from trace decay or reduced contextual overlap between learning and retrieval (as in SIMPLE and other distinctiveness theories), time-based forgetting in serial recall must, all else being equal, predict steeper primacy effects when the time course of retrieval is slowed.

Event-Based Models

Time-based theories stand in contrast to another class of models, which we call “event-based,” that assign little or no importance to the passage of time per se (e.g., Farrell & Lewandowsky, 2002; Henson, 1998; Lewandowsky, 1999; Lewandowsky & Murdock, 1989; Murdock, 1995; Nairne, 1990; Neath, 1999). That is, although these models acknowledge that many memory phenomena—such as forgetting—are observed over time, they consider the passage of time to be epiphenomenal, rather than causal. Instead, these models postulate a variety of other processes that cause forgetting without being informed by the passage of time. For example, several models assume the presence of output interference, such that the process of recalling each item interferes with the memory of yet-to-be-recalled information (e.g., Lewandowsky & Murdock, 1989). Other models postulate that earlier items are overwritten by new material during input (e.g., Nairne, 1990; Neath, 1999) and still others suggest that successive list items are encoded with progressively less strength (Farrell & Lewandowsky, 2002). For the purposes of this article, we do not differentiate between these processes because they all give rise to a common prediction—namely, that the

passage of time during output will not cause fanning of serial position curves. Instead, all things being equal, event-based theories predict that recall performance should be equivalent across all serial positions for all retrieval speeds.

Simulations: Comparing Architectures Within SIMPLE

We supplemented and formalized the preceding general statements about the predictions of the two classes of models by simulation. We used modifications of the SIMPLE temporal distinctiveness model to derive illustrative predictions, for the experiments that follow, from time-based and event-based forgetting mechanisms. For each mechanism, we explored the primacy effects that are predicted as a function of the pace of retrieval. Although SIMPLE, by default, implements time-based forgetting, its architecture is sufficiently flexible to permit a comparison of the two principal mechanisms within a common computational framework.

For the purposes of these simulations, a basic version of SIMPLE was used that did not incorporate mechanisms for extralist intrusions, did not incorporate response suppression, and ignored the presence of phonological or semantic similarity between items. All of these mechanisms are necessary for complete modeling of a range of paradigms (see Brown et al., 2002, for the complete specification of the model); however, they are not needed for (and indeed could potentially obscure) the theoretical predictions to be examined here. No rehearsal mechanism was included, and the predictions derived below therefore apply to experimental cases in which rehearsal is prevented or does not occur spontaneously.

SIMPLE makes three core assumptions: First, items are represented in terms of their position within a multi-dimensional psychological space (here, just two dimensions—positional and temporal—are used). Second, the similarity between any two items in memory is a reducing function of the distance separating them in that psychological space. Third, the probability of recalling a given item is inversely proportional to that item's summed similarity to all other potentially recallable items. We now specify each of these assumptions in more detail.

Encoding in a two-dimensional space. Representations in memory are organized along a primary dimension of temporal distance that reflects the (logarithmically transformed) elapsed time between study and retrieval. The remembered location along this temporal dimension provides a possible retrieval cue. The second dimension represents within-list position, coded by ordinal numbers (i.e., Position 1, Position 2, . . .). There is ample evidence that purely positional information is relevant in serial recall (Henson, 1999; Ng & Maybery, 2002), and the inclusion of a positional dimension is useful for the simulations we describe below. As an illustration of the interplay between those two dimensions, consider the representation of a two-item list after a 5-sec retention interval (assuming a 1-sec interitem interval at presentation). The two items would be in locations

$\{\log(6), 1\}$ and $\{\log(5), 2\}$, respectively, in the {time, position} space.

The relative importance of the two dimensions at retrieval is determined by the parameter wt , where wt is the attentional weight paid to the temporal dimension and $1-wt$ is the weight given to the positional dimension. More specifically,

$$D_{i,j} = wt|TD_i - TD_j| + (1-wt)|P_i - P_j|,$$

where $D_{i,j}$ is the psychological distance between stimulus i and stimulus j , TD_i is the (logarithmically transformed) temporal distance of stimulus i from the time of retrieval, and P_i is the ordinal position of item i . In intuitive terms, the psychological distance between two items depends on the sum of the distances separating the items on the temporal and positional dimensions. The attentional weight parameter (see, e.g., Nosofsky, 1992) can be thought of as stretching (or shrinking) the psychological space along the most (or least) important dimension. In the present case, if wt is unity, the simulations embody the time-based forgetting hypothesis (and for the two-item list above, the representation reduces to $\{\log(6)\}$ and $\{\log(5)\}$). Conversely, as wt approaches zero, forgetting is no longer time-based (and the representation of the two items becomes $\{1\}$ and $\{2\}$) but arises through an event-based process that is specified later in the context of the relevant simulations.

Similarity–distance metric. The second critical assumption, adopted from the categorization literature (see Brown et al., 2002, for extensive discussion), is that the similarity of any two items in memory is a reducing exponential function of the distance between them in psychological space. Thus

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

where $\eta_{i,j}$ is the similarity between items i and j , and D is the distance between them. The parameter c governs the rate at which similarity decreases with distance. This has the effect that items that are very close have a similarity approaching 1.0, whereas items that are more psychologically distant have a similarity that, in the extreme, approaches zero. In conjunction with the logarithmic transformation of time, this similarity metric gives rise to the distinctiveness ratios mentioned earlier when temporal distance is the only relevant dimension. That is, the similarity between any two items is the simple ratio between the elapsed times (the smaller divided by the larger) since study, raised to the power c .

Similarity determines recall. The third assumption is that the probability of recalling a given item $P(R_i)$ is inversely proportional to the summed similarity of that item to every other potentially recallable item. That is,

$$P(R_i) = \frac{1}{\sum_{k=1}^n (\eta_{i,k})},$$

where n is the number of items that could be recalled (i.e., the number of list items in the present simulations).

Omissions are assumed to arise from thresholding of low retrieval probabilities, as calculated by the preceding equation. Brown et al. (2002) implement this threshold mechanism with a sigmoid function such that if P is the recall probability from the preceding equation, it is transformed to become

$$\frac{1}{1 + e^{-s(P-t)}}$$

where t is the threshold and s (which gives the slope of the transforming function) can be interpreted as the noisiness of the threshold (see Brown et al., 2002, for a complete account). This mechanism ensures that any probability falling below the threshold will engender an omission. In addition, it will increase recall probabilities that are already high, and further reduce recall probabilities that are already low.

Model parameters and implementation. This basic model thus has four free parameters. One, c , governs the rate at which the psychological similarity of two items decreases as a function of the distance between them in psychological space. As c becomes larger, the confusability between items decreases more quickly as a function of their separation in psychological space, and memory performance increases. The second parameter, wt , specifies the amount of attention paid to the temporal dimension (at the expense of attention paid to the positional dimension). The third and fourth parameters, denoted s (for “slope”) and t (for “threshold”) relate to omissions as described above. These were held constant (at $s = 10$; $t = 0.6$) throughout the present simulations and do not materially affect the qualitative pattern of predictions, although larger values for these parameters tend to lead to steeper serial position curves overall.

All simulations implemented the presentation regime of the experiments reported below. Thus, it was assumed that items were shown at a rate of two per second, that recall of the first item commenced after a retention interval of 0.5 sec, and that items were recalled at a rate of 0.4 sec per item (“fast recall”), 0.8 sec per item (“medium recall”), or 1.6 sec per item (“slow recall”).

Predictions of time-based forgetting. We first examined the predictions of the model under the assumption that all attentional weight is given to the temporal dimension ($wt = 1$; equivalent to an exclusively time-based representation). For the reasons cited at the outset, it was expected that under these conditions the serial position curves would fan (with steeper primacy for slower recall rates). We also gradually reduced the attentional weight given to the temporal dimension ($wt = .75, .5$, and 0), which corresponds to progressively weaker versions of the time-based forgetting hypothesis, in the expectation that this manipulation would reduce the fanning of the serial position curves.

The parameter c was recalibrated for each setting of wt to achieve 60% correct performance overall in the “medium” speed condition. This ensured that the effect of varying wt could be observed without covarying overall

performance. (Without modifications to c , overall performance level would change as a function of wt for the theoretically uninteresting reason that the range of position values is greater than the range of values on the logarithmic temporal dimension.) The values of c were 10, 3.7, 2.3, and 1.32 (for $wt = 1, .75, .5$, and 0 , respectively).

The results of varying wt are shown in the leftmost column of Figure 1. It is evident that, as intuition suggests, there is considerable divergence of the serial position curves when most attention is paid to the temporal distance dimension. As attention to the temporal dimension is reduced and correspondingly shifted toward the positional dimension, two trends are evident. First, the predicted differences between retrieval speeds gradually diminish. Second, the extended primacy that is evident when wt is high gradually disappears and is replaced with a symmetrical serial position curve. This highlights the role of time-based forgetting in producing extended primacy.

In summary, to the extent that primacy exceeds recency in forward serial recall—which it invariably does in the data when list presentation is visual—time-based forgetting necessarily predicts fanning of the serial position curves if retrieval speed is manipulated. Time-based forgetting can handle the absence of fanning only at the expense of predicting unreasonably symmetrical serial position curves.

Event-based forgetting. Because SIMPLE is a temporal distinctiveness theory, no major conceptual decisions had to be made for the preceding simulations of time-based forgetting. For event-based forgetting, by contrast, several options existed for the simulations that paralleled the diversity among event-based theories (e.g., output interference vs. input interference; Lewandowsky & Murdock, 1989, vs. Nairne, 1990). We decided to model event-based forgetting through output interference. The attention weight for the temporal dimension, wt , was set to 0 , thus eliminating any time-based forgetting.

Output interference was implemented by assuming that the memory representations of to-be-recalled items become progressively less distinctive as recall proceeds. In intuitive terms, it is as if the recall of each item adds noise to the memory representation of yet-to-be-recalled items. In the simulations, we introduced a new parameter, o (for “output interference”) which reduced the value of c for the n th item recalled by multiplying $c \times o^{n-1}$. Thus with $o = 1$, there is no output interference; as o falls below 1 , there is increasing output interference in that the similarity gradients of each successively recalled item flatten (and thus overlap with those of competing items) ever more rapidly.

The right column of Figure 1 shows the predictions of the model with increasing amounts of output interference. The values of o were $1, .95, .9$, and $.8$, from top to bottom. As before, we adjusted c for each value of o to achieve about 60% correct recall in the “medium speed” condition. The values of c were 1.32, 1.52, 1.75, and 2.5 (for $o = 1, .95, .9$, and $.8$, respectively).

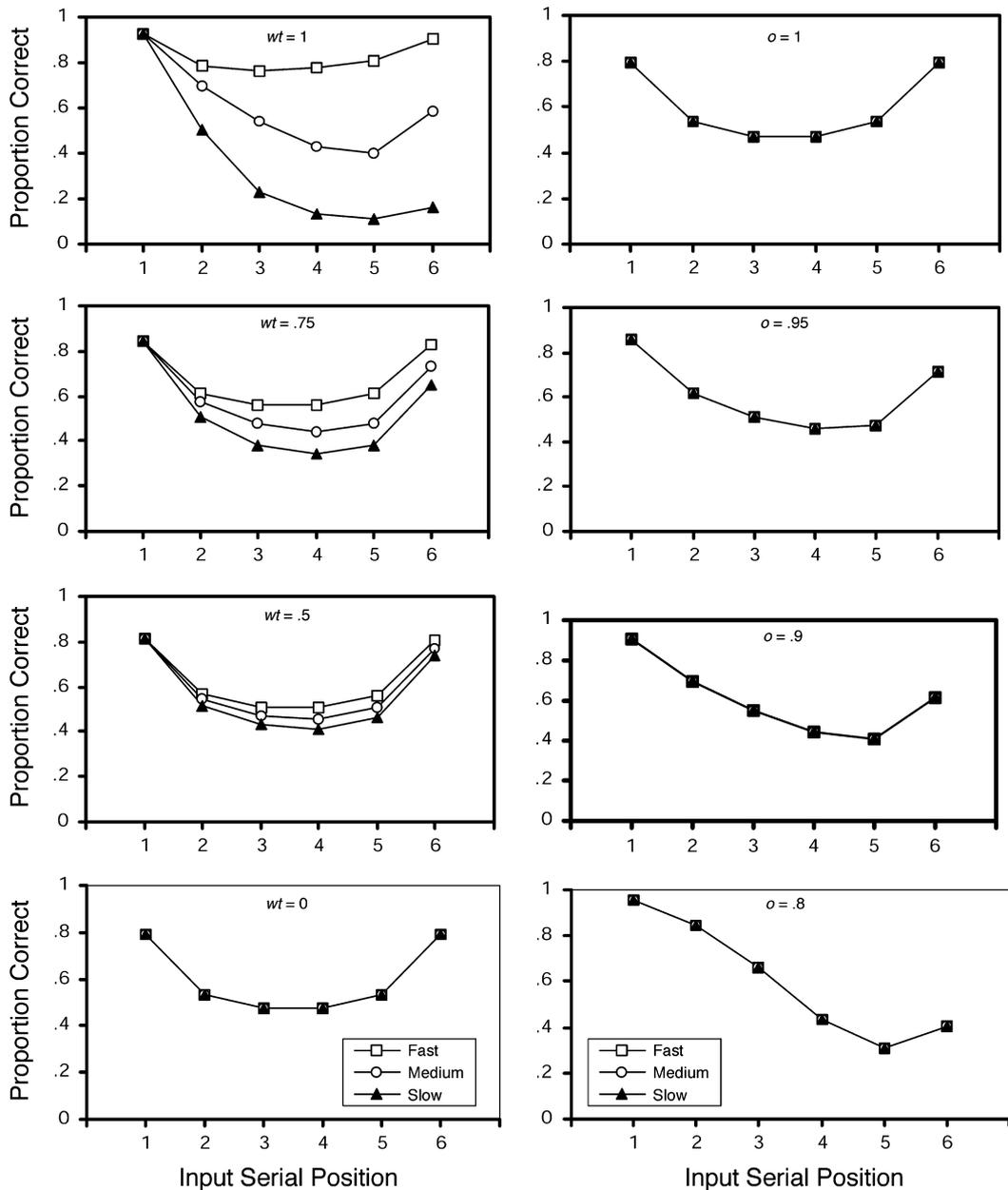


Figure 1. Predicted effects of retrieval time on the serial position curves for forward serial recall. The left column of panels shows predictions for a time-based model (with the attentional weight given to time shown in each panel). The right column shows predictions for an event-based model based on output interference (with the value of the output interference parameter shown in each panel).

It is evident from the figure that output interference had the expected effect. As output interference increases (from top to bottom across panels), so does the extent of primacy. However, because no time-based forgetting occurs, there is no effect of retrieval time, and the three serial position curves for the three recall speeds lie on top of one another. Thus, if forgetting occurs only due to output interference, recall speed has no effect on performance.

Because event-based forgetting can be instantiated in a variety of ways, in a final simulation we explored the alternative possibility that the positional encoding of later presented items becomes progressively less precise. This instantiates the assumption, which is at the heart of many models, that the quality of encoding decreases across serial positions (e.g., Brown et al., 2000; Henson, 1998; Lewandowsky, 1999; Lewandowsky & Li, 1994; Lewan-

dowsky & Murdock, 1989; Page & Norris, 1998). For these final simulations, we assumed that the positional dimension behaves, psychophysically, like other psychological dimensions by introducing a power-law compression. After compression, the position of item n was represented as n^p , where p is a parameter governing the amount of compression of the positional dimension (an alternative is to transform the positional dimension logarithmically, as is done with the temporal dimension, and this has a similar effect to power compression but does not readily permit parametric variation). When $p = 1$, no compression occurs. When p is less than 1, compression occurs with larger positional values being compressed to a greater extent than smaller positional values, thus rendering items toward the end of the list more psychologically similar. The role of p in this implementation thus paralleled the effect of o in the preceding simulations. The predictions associated with this alternative event-based implementation of forgetting turned out to be qualitatively indistinguishable from those provided by output interference and are thus not reported in the figure.

We conclude that event-based forgetting, whether instantiated by output interference or through a primacy gradient, yields predictions that are qualitatively different from those of time-based forgetting. According to event-based accounts, varying retrieval time while keeping all other variables constant should leave the serial position curves unchanged.

Events or Time?

The results of the simulations are clear. Time-based forgetting predicts substantial fanning of serial position curves as a function of retrieval speed, whereas event-based forgetting predicts no effect of speed.

Given these widely divergent predictions, and given the importance of this issue, it can be reasonably asked why the question has not yet been resolved empirically. Much research has addressed the role of time in memory, and many findings have been taken to support the role of time in forgetting. However, as we show next, many of the studies are open to alternative interpretations.

TIME AND FORGETTING: DATA

We distinguish three different classes of evidence that have been or could be taken as evidence for time-based forgetting during serial recall.

Time-Limited Forgetting Functions

Dosher (1999) and Dosher and Ma (1998) showed that as cumulative output time increased with increasing list length, perfect ordered list recall decreased (see also Wingfield & Byrnes, 1972, for an early demonstration of this effect). The strength of the relationship between output time and recall accuracy was underscored by the fact that it was identical for different types of material, for which accuracy and output time differed when each was

considered in isolation. However, it is difficult to determine the direction of causality from these results alone. It could be that items that are more difficult to recall also take longer to recall, and hence that the direction of causality runs from recall difficulty to output time, rather than the other way around (Crowder, 1976, in the discussion of the experiment by Wingfield & Byrnes, 1972, explicitly noted the correlational nature of this type of data).

A similar problem arises in the interpretation of the primacy effects seen in the serial recall performance of special populations, in which steeper serial position curves (reflecting relatively poorer recall of late-recalled items and/or increasing numbers of omissions in later output positions) may be seen in serial recall by elderly subjects (Maylor, Vousden, & Brown, 1999), children (McCormack, Brown, Vousden, & Henson, 2000), and patients with schizophrenia (Elvevåg, Weinberger, & Goldberg, 2001). In modeling developmental changes in serial recall at both ends of the life span, it has been necessary to assume additional forgetting during output in children (Brown, Vousden, McCormack, & Hulme, 1999) or elderly participants (Maylor et al., 1999), and this has typically been attributed to slower recall. However, the direction of causality again remains unclear (Brown et al., 2002): Poorer performance for late-recalled items may either cause, or be caused by, slower or more interference-susceptible recall.

Output Order Effects and Primacy in Serial Recall

The earlier simulations showed that both classes of forgetting mechanism can predict the extended primacy effect that is typical of forward serial recall. However, the accounts make divergent predictions concerning the effect of recall order on performance.

Evidence consistent with time-based forgetting comes from tasks that involve the serial recall of all list items in varying orders. For example, Cowan, Saults, Elliott, and Moreno (2002; see also Cowan, Saults, & Brown, 2004) deconfounded order of presentation and order of recall by requiring participants to recall nine-item lists in a "circular" manner, such that recall began with the first, fourth, or seventh presented item. In all cases, recall proceeded in forward order; in the latter two cases, the first three or the first six list items were reported after the initial sequence of forward recall (starting at Positions 4 or 7) was complete. The results of this procedure, which allow the effects of input and output position to be disentangled, showed that position of item recall was a major determinant of performance: Early recalled items tended to be recalled well, irrespectively of their input position. (Cowan et al., 2002, included additional examinations of modality effects and whole versus partial report that are of little relevance here; see also Beaman, 2002; Beaman & Morton, 2000.) These results provide strong evidence that *either* output interference *or* the passage of time during output is central to ordered serial

recall but, as Cowan et al. (2002) noted, the available results do not enable the roles of time and interference at output to be separated.

Item Length Effects

A third source of evidence for time-based forgetting concerns item length effects, which have often been ascribed to temporal factors during recall (e.g., Cowan et al., 1992). More specifically, the tendency for long words to be recalled less well than short words, when syllabic length is held constant (Baddeley, Thomson, & Buchanan, 1975), was initially assumed to reflect the operation of a time-based rehearsal process that could offset temporal decay. However, the time-based account of word length effects has become controversial. First, alternative models based on processes other than rehearsal have been developed (e.g., Brown & Hulme, 1995; Lewandowsky & Farrell, 2000; Neath & Nairne, 1995; see also Nairne, Neath, & Serra, 1997). Second, the appearance of a word length effect seems to depend upon the precise materials used. The effect is reliable for the materials used in early research (Baddeley et al., 1975; Cowan et al., 1992; Lovatt, Avons, & Masterson, 2000) but often does not appear when other materials are employed (Caplan, Rochon, & Waters, 1992; Lovatt et al., 2000; Service, 1998), leading to the suggestion that some kind of phonological complexity may be the relevant factor (e.g., Caplan et al., 1992; Caplan & Waters, 1994; Service, 1998). Likewise, recall-order effects, which show that performance is determined by the length of the words first recalled (Cowan et al., 1992; see also Cowan, Wood, Nugent, & Treisman, 1997) and which appeared to provide evidence that the passage of time at output is important, may be materials dependent (Lovatt, Avons, & Masterson, 2002). Overall, it appears inadvisable to put much faith in a result that provides only indirect support for time-based forgetting and does so only with certain words but not others.

In other studies, the same material was used in all conditions, and participants pronounced the words either quickly or slowly (e.g., Cowan et al., 2000; see also Cowan et al., 1997, and Service, 2000, for a critique). In the study by Cowan et al. (2000), participants were successfully trained to pronounce words at different speeds, and were asked to vocalize responses quickly or slowly during recall. The expected effect of pronunciation speed was obtained, with quick pronunciations associated with better recall. However, the pronunciation manipulation did not affect the total amount of time available for recall of each list item, which remained constant at 2.5 sec. Thus, participants had more time to rehearse or refresh the remaining list items after each response in the fast condition than in the slow condition (see Cowan, Nugent, & Elliott, 2000; Service, 2000). This differential rehearsal or memory search time could potentially explain the effects observed by Cowan, Nugent, and Elliott, in which case the data do not identify the passage of time as the causal variable underlying forgetting.

In summary, although the debate concerning the source of word length effects continues (e.g., Baddeley, Chincotta, Stafford, & Turk, 2002), we argue that the findings to date provide no compelling evidence for the causal role of time per se in forgetting.

Time and Forgetting

In overall summary of the three lines of evidence just reviewed, we suggest that, at present, no strong empirical conclusions are possible about the role of retrieval time in memory. Although at first glance there appears to be wide support for the assertion that recall accuracy declines with increasing retrieval time, closer inspection reveals that relevant studies either confounded retrieval time with the type of material (e.g., Baddeley et al., 1975), failed to control a possible contribution from rehearsal (e.g., Cowan, Nugent, & Elliott, 2000), or reported a correlation between two measures that left the direction of causality open (e.g., Doshier & Ma, 1998). We now turn to a recent precedent which controlled two of those factors—namely, type of material and direction of causality, before reporting two new experiments that additionally controlled rehearsal.

In a recent study by Duncan and Lewandowsky (2003, Experiment 1), time between encoding and retrieval was manipulated by extensively training participants, via corrective feedback, to recall items at speeds of either 400, 800, or 1,600 msec/item, following a constant presentation rate of 500 msec/item. Although Experiment 1 of Duncan and Lewandowsky was designed to address a different issue, when the data were analyzed with respect to retrieval speed, the resulting serial position curves showed no fanning. The absence of fanning is consistent with the predictions of event-based theories but contradicts the expectation of time-based forgetting. (Conrad & Hille, 1958, manipulated retrieval speed in a manner similar to that used by Duncan and Lewandowsky. However, interpretation of their results, which showed a slight advantage for fast retrieval over slow retrieval, is compromised in several ways. First, it is unclear whether participants in their study reported the items in forward order. Second, Conrad and Hille reported only probabilities of complete list recall, which further precludes examination of the serial position curves and hence assessment of fanning.²)

Because Duncan and Lewandowsky (2003) explicitly manipulated recall speed, the direction of causality in their study is clear. Moreover, Duncan and Lewandowsky used the same materials across all speed conditions, thus redressing a further difficulty associated with previous research. However, the study by Duncan and Lewandowsky did not control for rehearsal during output. That is, in the same way that participants in the experiment by Cowan, Nugent, and Elliott (2000) had more opportunity for rehearsal during fast than during slow retrieval, in the study by Duncan and Lewandowsky, rehearsal may also have differed between retrieval speeds, albeit in the op-

posite direction. Specifically, with the slower retrieval speeds, participants may have had more opportunities to rehearse items in between retrievals, thus obscuring any effects of time. (Rehearsal was also left uncontrolled in the study by Conrad & Hille, 1958.) The present experiments, then, were designed to control rehearsal at retrieval while manipulating recall speed with the technique introduced by Duncan and Lewandowsky.

EXPERIMENT 1

Following Duncan and Lewandowsky (2003, Experiment 1), participants were trained to recall lists at different speeds. Unlike in the earlier experiment, participants sometimes engaged in articulatory suppression (AS) during recall to prevent rehearsal.

If time is a causal factor in recall performance as time-based theories predict, there should be a significant interaction between recall speed and serial position when rehearsal has been eliminated by articulation (i.e., the serial position curves should show fanning). To the extent that rehearsal is absent even without articulation, a similar fanning could be observed without articulatory suppression. If, on the other hand, time is not a causal factor, no interaction should be present (i.e., the differences between serial position curves for the different speeds, if any, should not vary across serial positions).

Method

Overview and Design. Participants were trained to recall lists of six items at 400, 800, or 1,600 msec/item. During the first experimental session, presentation and recall of lists was unaccompanied by AS. In the second session, participants were required to engage in AS during recall by speaking a suppressor word out loud. Order of sessions was not counterbalanced in order to ensure that the first session would represent a replication of Duncan and Lewandowsky (2003, Experiment 1). Thus, the experiment represented a $2 \times 3 \times 6$ within-participants design with suppression (quiet vs. AS), recall speed (400, 800, or 1,600 msec/item), and serial position (Positions 1–6) as factors.

Participants. Eleven members of the University of Western Australia campus community participated voluntarily in two 1-h sessions. Participants were remunerated at the rate of A\$10 per session.

Apparatus and Materials. All lists contained the six letters H, J, M, Q, R, and V in a different random order for each trial. For each speed condition, 40 lists were constructed, resulting in a total of 120 lists per session. Recall speeds were alternated randomly across trials for each participant in each session.

A PC was used to present each list and, via keyboard entry, record both the item recalled and the interresponse times (IRTs). Participants sat approximately 45 cm from the screen and used the keyboard to record each recalled response. In the second session, a vocalization of the suppressor was recorded to ensure compliance with AS instructions (compliance was subsequently verified by inspection of audio recordings).

Procedure. Each trial started with the signal to “get ready,” followed by a 1,000-msec blank interval. Each list item was then presented individually in the center of the screen for 400 msec, followed by a blank interval for 100 msec. After list presentation, there was a blank interval of 500 msec, followed by a word in the center of the screen that cued recall speed. The word was FAST for 400 msec/item, MEDIUM for 800 msec/item, or SLOW for 1,600 msec/item. In addition, an underscore character appeared below and to the

left of the visual cue. As participants recalled each item, their response replaced the underscore character, which then moved to the right. After participants had recalled six items, they were given feedback regarding their actual speed of recall. After reviewing their feedback, participants pressed a key to initiate the next trial.

Feedback was provided by computing the deviation of each recorded IRT from the required IRT (e.g., for a speed of 400 msec/item, the required IRT was 400 msec). If the recorded IRT was outside a predetermined criterion, that item was flagged as either too fast (recorded IRT fell below the criterion) or too slow (recorded IRT fell above the criterion). Pilot experimentation determined optimal settings for the IRT criteria, which were ± 50 msec for the 400 msec/item speed, ± 70 msec for 800 msec/item, and ± 120 msec for 1,600 msec/item. These criteria corresponded roughly to ± 1 standard deviation of IRTs observed in the pilot sessions. If two or more items fell outside the criterion bounds of the required IRT, participants were informed that their recall was “too fast” (or “too slow”). If there was a mixture of two or more responses that were too fast or too slow, participants were informed that their recall was “uneven.” If all recorded IRTs were within the criterion bounds, participants were told “speed ok.”

The number of trials in which a participant’s speed did not match the required speed was also recorded (i.e., cases in which “too slow,” “too fast” or “uneven” feedback was given). If a participant failed to match three successive trials of a particular speed, an auditory signal was presented in addition to the usual feedback that consisted of a series of tones played at the required speed. The auditory signal was composed of a series of 400-Hz tones of 100 msec duration each, spaced at the IRT representing the beat of the required speed for that trial. The auditory signal lasted for 3.2 sec.

Participants were instructed to report the list items in forward serial order on each trial. For Session 2, participants were instructed to speak the suppressor (the word “super”) during retrieval of the list at the rate of the fastest speed (400 msec/item).

Experimental trials of Session 1 were preceded by six blocks of practice trials. Each block consisted of 3 trials at the same speed, thus yielding a total of 18 practice trials (6 at each speed). In the practice (but not experimental) trials, participants were cued about retrieval speed before list presentation, and the cue was accompanied by the same auditory signal used during corrective feedback.

Results and Discussion

Recall speed manipulation. Figure 2 shows frequency histograms of the observed IRTs for quiet (panel A) and AS (panel B) recall. Frequencies were computed for each speed condition by aggregating across serial positions and participants. Vertical lines mark the target IRT for each speed. It is clear from both panels that participants learned to recall the lists at different speeds irrespective of the presence or absence of AS.

Although there is some overlap between distributions, they all peak at or very near the true IRT, especially when AS was present. We therefore conclude that the recall speed manipulation was successful.

Serial position curves. To assess recall performance, the total number of trials in each session in which an item was recalled in its original input position was computed for each participant. These sums were converted to log odds (cf. Henson, 1998) and then entered into a $2 \times 3 \times 6$ within-participants analysis of variance (ANOVA), with the factors representing presence of suppression (quiet vs. AS), recall speed (400, 800, or 1,600 msec/item), and serial position (1–6). If people in the earlier studies by Duncan and Lewandowsky (2003) were rehearsing dur-

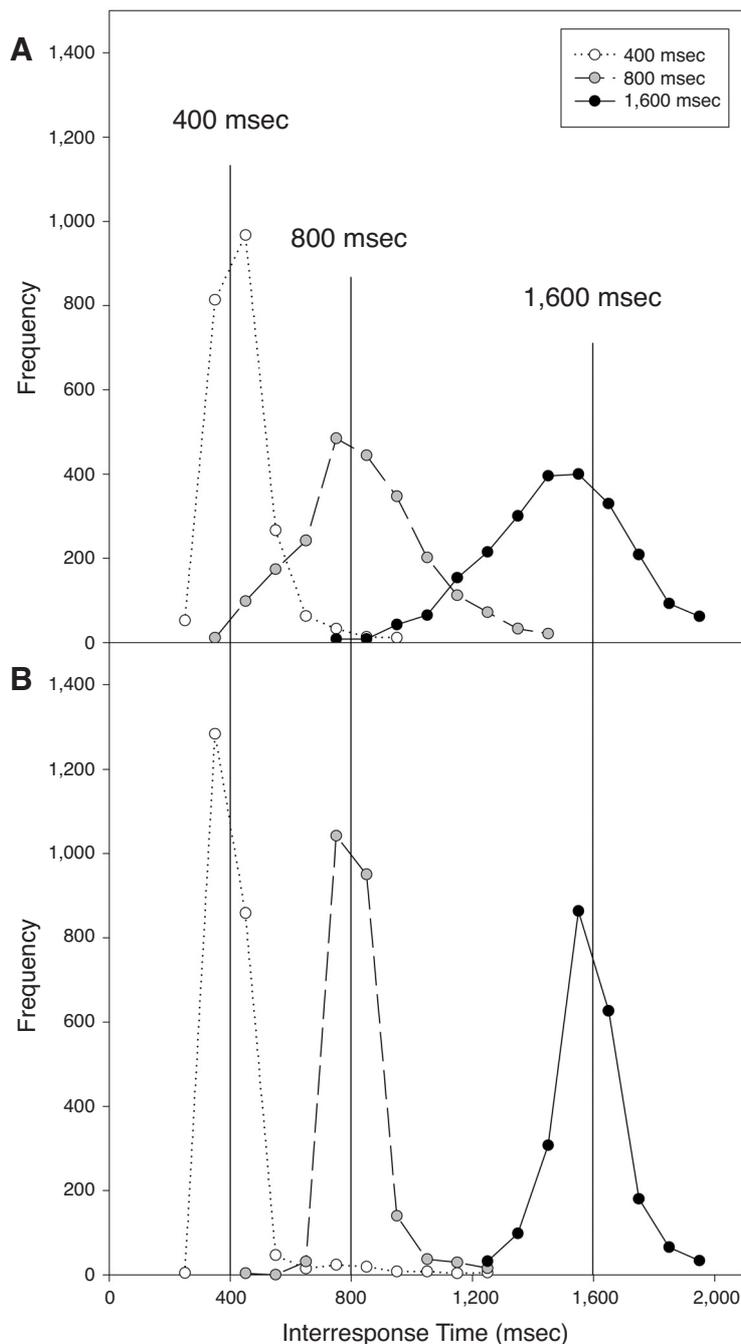


Figure 2. Frequency histograms of interresponse times for three speed conditions during quiet recall (A) and recall with articulatory suppression (B) in Experiment 1.

ing retrieval, thus obscuring any effects of recall speed, time-based theories would predict a significant interaction between serial position and speed when AS was present, but not when recall was quiet. This pattern of interactions, in turn, would be reflected by a significant overarching three-way interaction. By contrast, event-based theories expect speed and serial position not to

enter into any interactions together because fanning of serial position curves should not occur. The absence of fanning is expected, whether or not people can rehearse during recall.

The ANOVA revealed significant effects of speed [$F(2,20) = 6.11, MS_e = 0.41, p < .05$], serial position [$F(5,50) = 46.11, MS_e = 0.91, p < .001$], and a signifi-

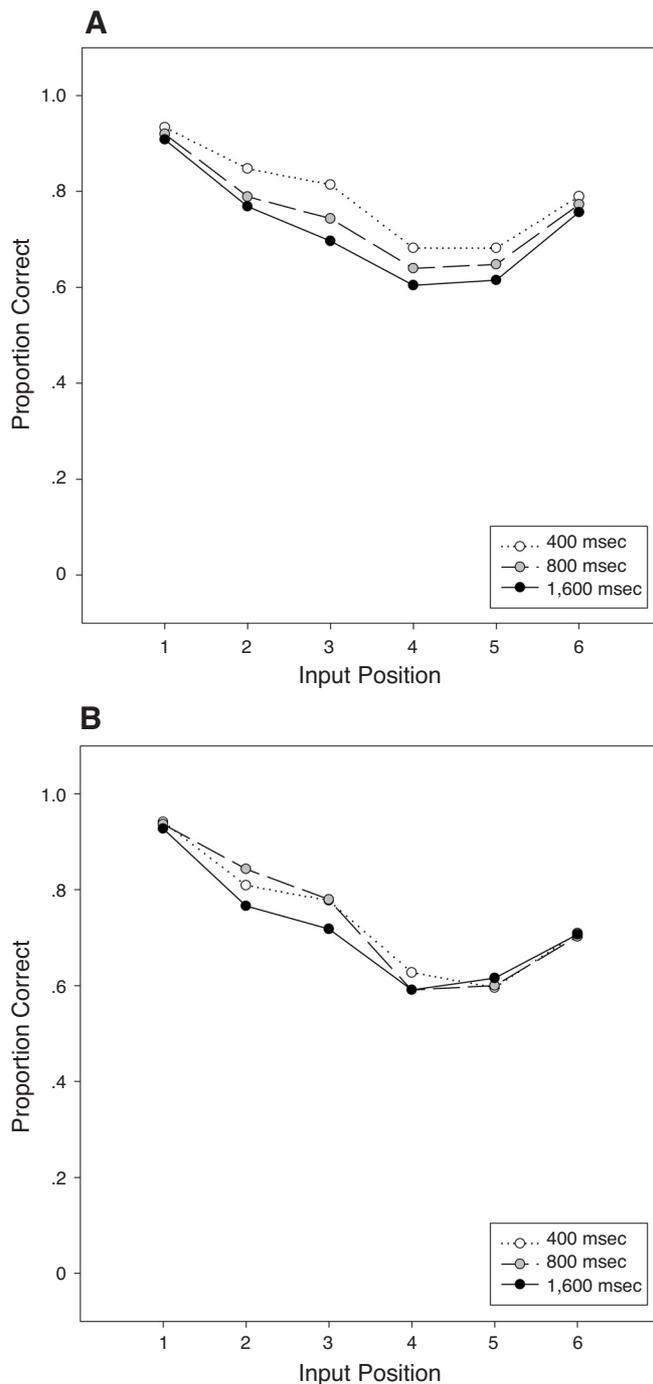


Figure 3. (A) Proportion correct in position recall for quiet retrieval for the 400-msec/item (white circles), 800-msec/item (gray circles), and 1,600-msec/item (black circles) speed conditions. (B) Proportion correct in position recall under articulatory suppression for the 400-msec/item (white circles), 800-msec/item (gray circles), and 1,600-msec/item (black circles) speed conditions in Experiment 1.

cant serial position \times suppression interaction [$F(5,50) = 2.68$, $MS_e = 0.32$, $p < .05$]. However, the three-way interaction was nonsignificant [$F(10,100) = 0.62$, $MS_e = 0.20$, $p < .60$], and speed also did not interact

with serial position [$F(10,100) = 1.15$]. The serial position curves in Figure 3A (quiet) and Figure 3B (AS) clarify that there was no hint of interaction involving recall speed in either condition. Clearly, retrieval time did not

have the effect on performance that is expected by time-based theories.

Nonetheless, the main effect of retrieval speed was significant, and at first glance may appear to favor time-based theories while compromising event-based views. However, that interpretation does not withstand scrutiny. First, and most important, if time caused forgetting, as we showed by simulation earlier, this would necessarily be reflected in fanning (and hence an interaction with serial position) and not a main effect of speed. In fact, a time-based theory would have difficulty accounting for a main effect of retrieval speed. Second, the main effect of time can be explained by revisiting the observed IRTs in Figure 2. It

is clear that people had greater difficulty pacing themselves at the slower speeds. It is therefore quite likely that they had to devote more attention to the timing task at slower speeds, and this additional load may have lowered recall performance overall in the slower conditions.

Two other aspects of the data merit consideration. First, the serial position \times suppression interaction reflects the fact that articulation had little or no effect for the first item recalled. This is readily explained in that very little time elapsed between the signal to commence suppression and recall of the first item. Hence, it comes as little surprise that the first item (or indeed the first two) was less affected by articulation overall. Finally, the

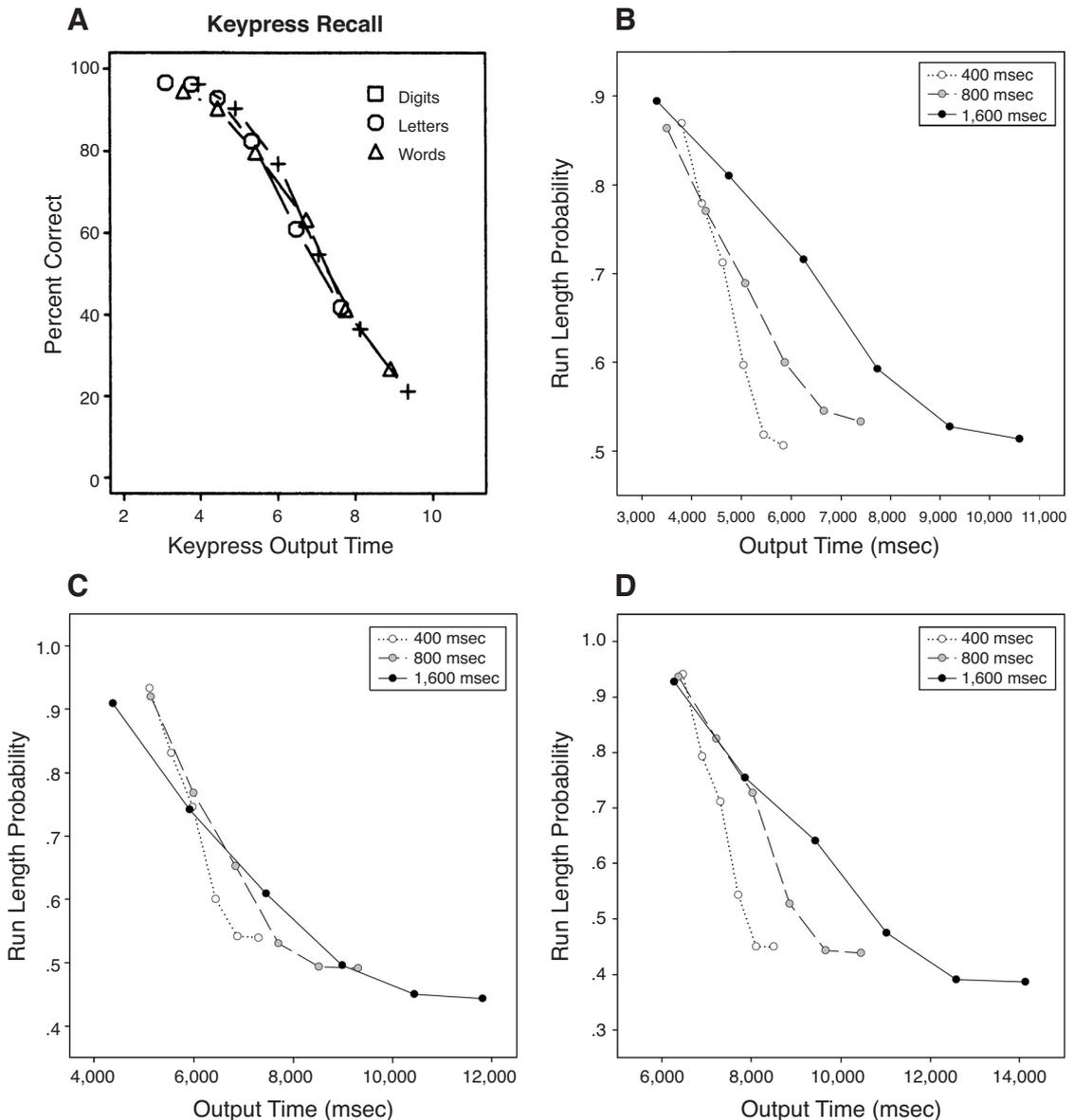


Figure 4. Run length probability for each speed as a function of output time (cumulative over output position). Panel A shows data from Doshier and Ma (1998, Figure 5B). Panel B shows data from Duncan and Levandowsky (2003, Experiment 1). Panel C, quiet condition from Experiment 1. Panel D, articulatory suppression condition from Experiment 1.

preceding interpretation of the interaction also helps explain the absence of a main effect of suppression [$F(1,10) < 1$], which one might expect at first glance because articulation has a dramatic effect upon performance when required during list presentation (e.g., Lewandowsky & Brown, in press). A further reason for the absence of the main effect here can be found in other work that has shown that dividing attention is far less consequential at retrieval than it is at encoding (e.g., Craik, Govoni, Naveh-Benjamin, & Anderson, 1996).

Effects of practice or proactive interference. We next explored whether people's performance changed across the first experimental session. A performance change might be expected in either of two ways, each with an opposing outcome: First, one might expect people to improve with practice. Second, one might expect performance to drop off as proactive interference (PI) builds up across successive lists. The buildup of PI might be particularly relevant in Experiment 1 because the same items were shown on all trials.

We thus compared performance on the first 10 trials at each speed to the last 10 during the first (quiet) session using a speed \times serial position \times trial within-participants ANOVA. The effect of trial (first 10 vs. last 10 at each speed) was significant [$F(1,10) = 8.32, MS_e = 2.08, p < .02$], but trial did not enter into any interaction with speed (largest $F = .34$). The main effect reflected the fact that people performed better late in the session, thus showing an improvement with practice and confirming that PI played little if any role in Experiment 1.

Memory performance over time. We now provide another perspective on the relationship between retrieval time and accuracy by examining cumulative accuracy as a function of elapsed output time. In the studies by Doshier and colleagues mentioned earlier (Doshier, 1999; Doshier & Ma, 1998), it was found that the probability of correctly recalling an entire list decreased as total recall time for the lists increased (manipulated by increasing list length). Panel A in Figure 4 shows representative results from Doshier and Ma (1998, their Figure 5B). The parameters in panel A correspond to different materials (letters, words, and digits) and, for a given parameter, each data point represents accuracy and output time for a different list length. Because type of material did not affect recall, whereas performance uniformly declined with output time, Doshier argued that "... temporal limits on immediate memory are ... temporal limits on recall output" (Doshier, 1999, p. 277). However, because output time was not experimentally manipulated but covaried with list length, the resulting unclear direction of causality implies that their results cannot distinguish between time-based and event-based theories. As Doshier and Ma themselves note, a key issue with their results is "whether output times are the cause or the consequence of forgetting. Is performance worse for words because they are more difficult to remember and therefore take longer to output during recall? Or is performance worse for words because they take longer to output during recall and, hence, more is forgotten?" (p. 329).

In the present experiment, by contrast, output time was manipulated independently of list length, thus permitting identification of the direction of causality. The present data can be presented in a way analogous to Doshier's output time functions (panel A in Figure 4), by plotting run length probability as a function of cumulative output time. Cumulative output time for a given item was obtained by adding the IRT of that position to the summed IRT of all previous output positions (e.g., the cumulative output time for the third recalled item was the sum of IRTs for output Positions 1–3). Run length probability was determined by scoring correct-in-position recall up to the first error (where an error is defined as an intrusion or transposition) for each output position. For example, a run length probability of .2 for Output Position 4 implies that on 20% of all trials the first four responses were correct. This measure is nearly identical to that used by Doshier and colleagues except that in our case, the participants determine "list length" by the position of their first error, whereas in Doshier's experiments list length was manipulated by the experimenter.

Panel B in Figure 4 shows the results of Duncan and Lewandowsky (2003, Experiment 1), and panels C and D show the equivalent data for the quiet and AS conditions, respectively, of the present experiment. Within each panel, the three parameters represent the three retrieval speeds. In replication of Doshier (1999) and Doshier and Ma (1998), panels B–D show a clear effect of time for each recall speed. However, the curves for the different retrieval speeds diverge considerably, implying that recall was not solely determined by the passage of time: In all panels, a given level of accuracy was associated with several output times. For example, in panel D, a constant run-length probability of .5 was obtained after nearly 8 sec of recall (for 400 msec/item), 9 sec (800 msec/item), or nearly 11 sec (1,600 msec/item). This result cannot be accommodated by time-based models which predict that a single function should relate accuracy to total output time across all speeds. By implication, the results also identify the earlier conclusions by Doshier and Ma as being problematic. Nonetheless, in some cases (e.g., points 3–6 in panel D), a slight effect of output time appears to be present because those points are lowest in the slowest retrieval condition.

Model-based analysis. On the basis of visual inspection alone, the results of Experiment 1 were at odds with the predictions of the time-based forgetting hypothesis (left column of Figure 1). However, this does not preclude the possibility that a model that includes time-based forgetting may provide a (slightly) better account of the data than would a pure event-based model. We therefore fitted a version of SIMPLE to the results of Experiment 1 to examine whether the fit would be significantly improved by including attention to the temporal dimension during retrieval in addition to output interference.

We used a constrained version of the model in which no omissions were possible, thus removing two free parameters and reflecting the procedure of Experiment 1 as closely as possible (although no omissions were possible

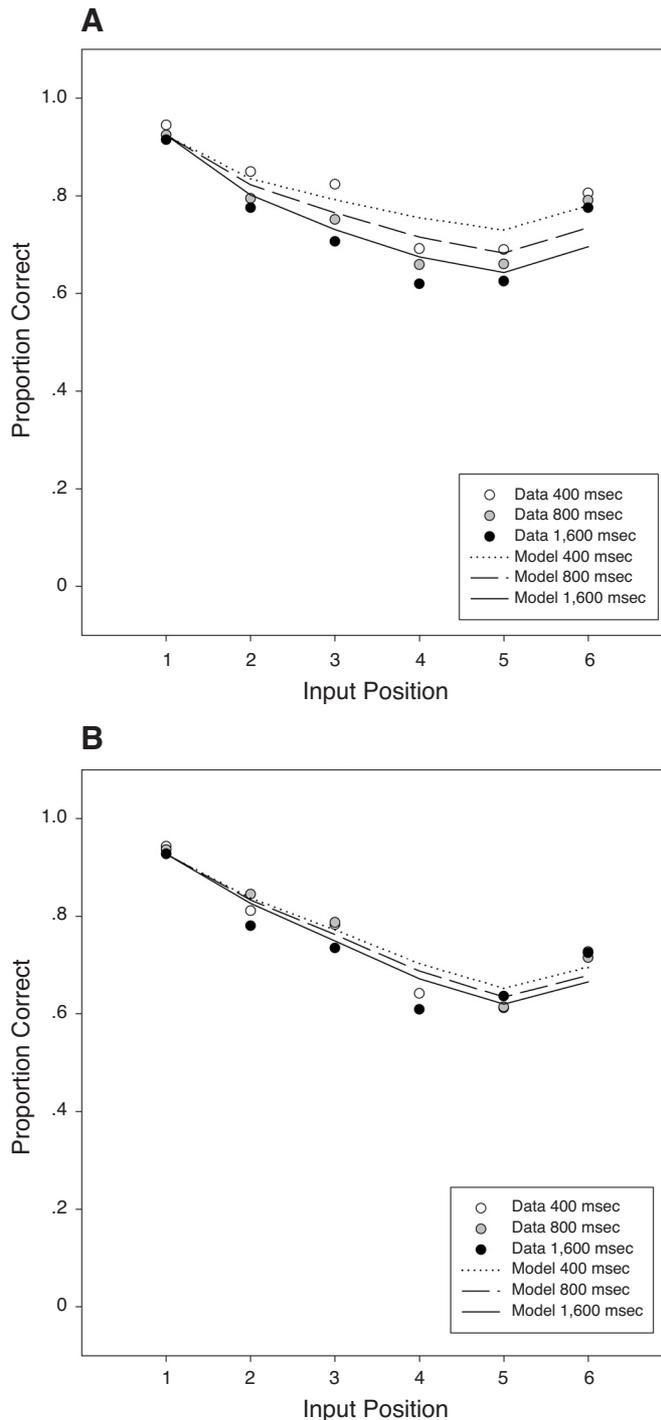


Figure 5. SIMPLE's predicted serial position curves (lines) and obtained data (symbols) for all speed conditions in Experiment 1. Panel A shows quiet recall, and Panel B shows recall during articulatory suppression.

in Experiment 1, participants might occasionally have intruded extralist items or committed repetition errors, but it appeared unnecessary to represent these rare events in the model). There were thus three free parameters: The

generalization parameter c , the attention paid to the temporal dimension, wt , and the amount of output inhibition, o . The modeling logic was as follows. We fit a general version of the model, and two restricted versions, to data

from individual participants. In the general version, all three parameters were free to vary for each participant. In one of the restricted versions, wt was set to zero, and so whatever forgetting occurred was entirely due to output interference (with o freely estimated). In the second restricted version, o was set to 1 and so only time could contribute to forgetting. In this version of the model, wt was freely estimated, thus allowing the model to consider position as well as time in representing the list, without, however, allowing any forgetting other than through the passage of time.

Each of the restricted models could then be compared with the general model using residual sums of squares (Borowiak, 1989). More specifically, comparisons between general and restricted models were calculated using a likelihood ratio test as follows:

$$\chi^2 = -2 \ln \left(\frac{\text{RSS}_{\text{GEN}}}{\text{RSS}_{\text{RES}}} \right)^{\frac{n}{2}},$$

where n is the number of data points, RSS_{GEN} is the residual sum of squares for the general model (both wt and o included) and RSS_{RES} is the residual sum of squares for the restricted model (either $wt = 0$ or $o = 1$).³ To accommodate the multiple model comparisons being undertaken ($N > 40$ across both experiments), we set a conservative ($p < .001$) criterion for the model fit comparisons.

Conditionalized response proportions were used for both predictions and data. That is, all entries in the transposition matrix, where rows represent input positions and columns represent output positions, were divided by row totals.⁴ Parameter values that minimized the residual sum of squares for the conditionalized transposition matrix were obtained by simulation. One consequence of using the complete transposition matrix to estimate parameters is that the R^2 values can be expected to be high for all versions of the model. The transposition matrix contains a large number of near-zero values, and all models can be expected to account for those quite readily (i.e., no model would expect an unrealistically large number of, say, transpositions across four positions). It follows that absolute R^2 values must be interpreted with great caution. Crucially, however, model comparison is unaffected by the absolute values of R^2 .

The fit of the unrestricted model (allowing effects of both time and output interference) to the data of the quiet condition is shown in panel A of Figure 5. The figure shows the mean of the model's fits to individual participant data. The median R^2 value for the unrestricted model was .965; this was unchanged when $wt = 0$ (i.e., when only output interference was included in the model) but was reduced to .947 when $o = 1$ (i.e., when only time, and not output interference, contributed to forgetting). For the unrestricted model, the mean parameter values were 6.45 (c), .53 (wt), and .90 (o), with the corresponding standard deviations being 3.43, .37, and .07, respectively. Note that the absolute value of wt depends on the

units of temporal encoding in the model; all times were coded in seconds.

For just 1 out of the 10 participants, the fit worsened appreciably when the temporal dimension was excluded from the model [$\chi^2(1) = 12.6$], although the effect approached the conservative significance criterion ($p < .001$) for 2 other participants. Conversely, the fit was significantly worse for 5 of the 10 participants when output interference was excluded [lowest $\chi^2(1) = 10.7$].

A similar analysis was carried out for the suppression condition. Panel B of Figure 5 shows the mean fit of the unrestricted model. When applied to the data sets from individual participants, median R^2 values of .962, .962, and .946 were obtained for the full model, the no-time model, and the no-interference model, respectively. A significant reduction in fit was observed for only 1 participant when time-based forgetting was excluded from the model [$\chi^2(1) = 11.6$], whereas fit was reduced for 6 participants when output interference was excluded [lowest $\chi^2(1) = 19.9$]. The mean parameter values for the unrestricted model were 4.76 (c), .24 (wt), and .84 (o), with the corresponding standard deviations being 3.55, .34, and .10, respectively.

The results of the model-based analysis are thus entirely consistent with the preceding conventional analyses. The data strongly support the hypothesis that event-based output interference, rather than time-based forgetting, underpins the serial position curves obtained in Experiment 1.

Summary and Limitations

Experiment 1 demonstrated clearly that output time is of little relevance in serial recall. Although participants demonstrably recalled the lists at different speeds, this manipulation did not engender any fanning of the serial position curves. The results could not have been due to rehearsal because the data remained unchanged if rehearsal during recall was prevented by AS.

One possible limitation of Experiment 1 was the use of keyboard entry during recall. Typed recall may have conceivably permitted participants to recode the list as a "motor program," perhaps obtained through visual inspection of the keyboard prior to recall. Such a motor program may have been immune to decay over time and not subject to interference by AS. Although this possibility does not negate or alter our principal conclusions, it may limit their generality. Experiment 2 therefore used spoken rather than typed recall to extend the generality of our findings. Because the absence of an effect of time is most informative when rehearsal is prevented, AS was used throughout Experiment 2.

EXPERIMENT 2

Experiment 2 differed from the first study in two main respects: First, participants recalled lists verbally, rather than by typing responses. Second, recall speed was manipulated by having participants speak the suppressor

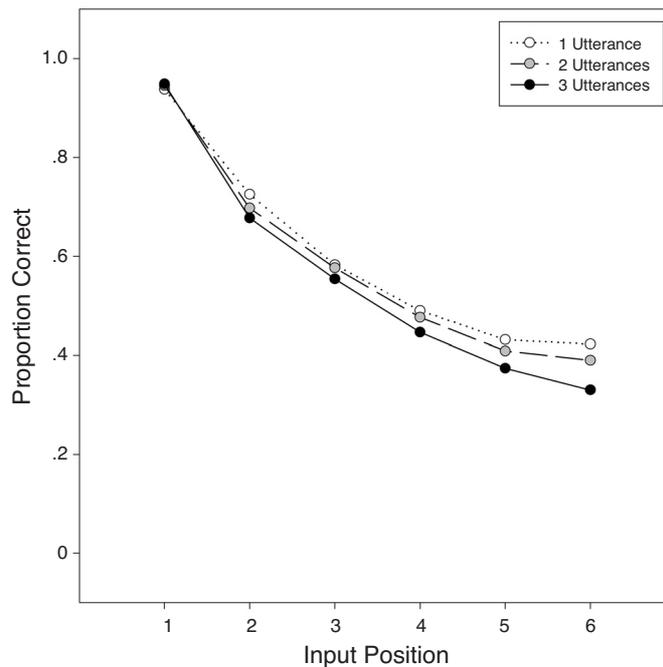


Figure 6. Proportion correct-in-position recall for the three speed conditions—one utterance (white circles), two utterances (gray circles), and three utterances (black circles)—in Experiment 2.

once, twice, or three times in between retrievals to instantiate the fast, medium, and slow conditions, respectively. This permitted manipulation of retrieval speeds without requiring training of participants and also rendered it more likely that pacing was equally difficult at all speeds.

Method

Overview and Design. As in Experiment 1, there were two experimental sessions, but both sessions were identical and involved AS during retrieval. Thus, Experiment 2 used a 3×6 within-participants design with speed (1 utterance, 2 utterances, or 3 utterances), and serial position (Positions 1–6) as the only factors.

Participants. Ten members of the campus community participated voluntarily in two 1-h sessions. Participants were remunerated at the rate of A\$10 per session.

Materials. Materials were the same as in Experiment 1. For each session, 40 lists were constructed for each speed by randomizing the order of the six list items, creating a total of 120 lists per session. Order of recall speed was again randomized across trials for each session.

The experimenter remained present to record the spoken recall and to ensure that participants complied with the AS instructions. For half of the participants, audio recordings were made as in Experiment 1 for later analysis of total recall time.

Procedure. The procedure was the same as in Experiment 1, except that the speed cue used was “1 trial” for one utterance, “2 trial” for two utterances, and “3 trial” for three utterances. There was also no speed feedback at the end of recall, and recalled items were not displayed on the screen. Instead, the experimenter recorded responses, and participants pressed a key on the keyboard to initiate each trial. Although participants were required to press a key on each trial, the keyboard itself was positioned such that participants could not actually view any of the keys.

Participants were instructed to report the list items in forward serial order and to repeat the suppressor (“super”) between recalls as dictated by the cue. As in Experiment 1, the experiment proper was preceded by practice trials that were identical to experimental trials. For half of the participants, responses were digitally recorded for later analysis of recall time.

Results and Discussion

Recall speed manipulation. An algorithm that measured both spike frequency and spike amplitude was used to detect voice onset of the first recalled item and voice offset of the last recalled item. Parameters for each criterion (frequency and amplitude) were set for each participant by sampling baseline noise from the participant’s recorded output. The total time difference between recall of the first item and the last item was then computed for each trial and aggregated across trials at each level of speed. Analysis of mean total recall times (6.46, 7.91, and 9.93 sec for 1, 2, and 3 utterances, respectively) by one-way ANOVA with recall speed as the single factor revealed a significant effect [$F(2,4) = 51.92$, $MS_e = 0.03$, $p < .001$]. This outcome confirms that the speed manipulation had the intended effect.

Serial position curves. Correct-in-position serial position curves were computed as for Experiment 1 and were analyzed using a 3×6 within-participants ANOVA. As before, time-based theories would predict a significant two-way interaction reflecting the fanning of serial position curves, whereas event-based theories predict the absence of the two-way interaction. The analysis produced a significant effect of speed [$F(2,18) = 3.62$,

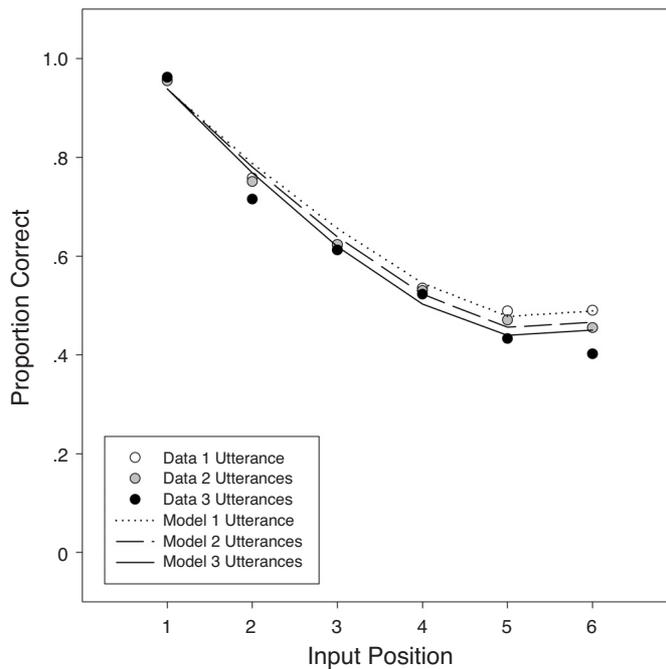


Figure 7. SIMPLE's predicted serial position curves (lines) and obtained data (symbols) for all speed conditions in Experiment 2.

$MS_e = 0.21, p < .05$] and a significant effect of serial position [$F(5,45) = 77.57, MS_e = 0.85, p < .001$]. The interaction of the two variables, however, was not significant [$F(10,90) = 0.83, MS_e = 0.08$].

The data are shown in Figure 6. It is clear from the figure that the effect of speed was small in magnitude and did not vary appreciably across serial positions. However, the figure reveals a small degree of fanning, and one might therefore wonder if the interaction would have been significant had the experiment had greater statistical power. The model-based analysis addresses this possibility.

Model-based analysis. SIMPLE was applied to the results of Experiment 2 in exactly the same way as in Experiment 1. The mean fit of the complete model (i.e., including time and output interference) to the individual participant data is shown in Figure 7.

A median R^2 of .948 was obtained for the unrestricted model. This reduced to .947 when $wt = 0$ (i.e., when only output interference was included) and to .880 when $o = 1$ (i.e., when only time, and not output interference, contributed to forgetting). Model comparison revealed no significant reduction in fit for any participant when the temporal dimension was excluded (when the conservative $p < .001$ criterion was used). However, when the model excluded output interference, the data from every participant were less well fit [lowest $\chi^2(1) = 14.2$]. The mean parameter values for the unrestricted model were 5.36 (c), .40 (wt), and .73 (o), with the corresponding standard deviations being 3.04, .34, and .09, respectively. For all participants, item-based output interfer-

ence, rather than time-based forgetting, thus provides the best explanation of the data from Experiment 2.

The model-based analysis thus converges with the results of the conventional statistical analyses, with one important additional consideration: The model-based analysis is immune to the argument that the slight fanning observed in Figure 7 might have failed to reach significance merely because the experiment lacked statistical power due to insufficient participant numbers. Even if the slight fanning had been statistically significant (but otherwise unchanged) in the conventional analysis, the effect size, and hence the conclusions from the model-based analysis, would have remained the same.

Summary

Experiment 2 replicated and extended the generality of the findings of Experiment 1 and of the earlier study by Duncan and Lewandowsky (2003). In all cases, retrieval time was manipulated without changing materials and, even when rehearsal was prevented by AS, no effect of retrieval time on accuracy was obtained. Experiment 2 confirmed that this result did not reflect the recoding into a "motor program" that might have been associated with the keyboard recall used in Experiment 1.

Experiment 2 also replicated the finding of Experiment 1, that AS by itself caused little additional interference. If fanning of the serial position curves had been obtained under AS in either experiment, one possible interpretation would have been that interference induced by the suppression itself (rather than the passage of time) was

responsible for the effect. The absence of any such effect confirms that AS, while surely preventing rehearsal, did so without introducing additional interference.

GENERAL DISCUSSION

Relationship to Previous Findings

The present results disconfirmed the predictions of time-based forgetting and, by the same token, endorsed an event-based mechanism instead. This conclusion was based not only on conventional interpretation and statistical analysis but also on quantitative modeling. This outcome is at odds with many other findings previously reported in the literature, and it is therefore important to restate and examine the differences between our experiments and previous methodologies. We identify five such differences that underscore the contribution of our studies.

First, retrieval time in both experiments (and in Experiment 1 of Duncan & Lewandowsky, 2003; see also Conrad & Hille, 1958) was under direct experimental control. Thus, unlike in some earlier studies that observed a correlation between recall accuracy and retrieval time (e.g., Doshier, 1999; Doshier & Ma, 1998; Wingfield & Byrnes, 1972), the direction of causality in our experiments is clear: Any effect of time on the fanning of serial position curves, had it been present, would have been causal.

Second, our manipulation of retrieval speed was not confounded with other variables. In particular, the stimulus materials were kept constant for all speeds. Lest one think that this control is a trivial experimental feature, it must be borne in mind that demonstrations of the word length effect, which previously has been taken to support time-based forgetting, are notoriously sensitive to the particular words being used (e.g., Lovatt et al., 2000).

Third, the use of articulatory suppression in both experiments, and the fact that articulation did not alter the role of retrieval speed, ensures that differential rehearsal could not have obscured an effect of time. This is again in contrast to previous studies in which differential rehearsal may have either caused an effect of speed (e.g., Cowan et al., 2000) or may have obscured the presence of such an effect (Conrad & Hille, 1958; Duncan & Lewandowsky, 2003; see also Waugh & Norman, 1965).

Fourth, the absence of time-based forgetting was not tied to a particular way in which people recalled the list. Specifically, it could not have been a consequence of recoding the list into a "motor program" before initiating keyboard recall in Experiment 1. The results of Experiment 2 confirmed that time-based forgetting is also absent with verbal recall of the list. However, we do note that both experiments, as well as the study by Duncan and Lewandowsky (2003), involved forward serial recall. We therefore cannot preclude the possibility that different results may be obtained with, say, probed recall or backward recall (cf. Li & Lewandowsky, 1993, 1995).

Fifth, the use of a model-based analysis in both experiments reduces the likelihood that our conclusions are

based on erroneous acceptance of the null hypothesis. Although it is in principle possible that both experiments lacked enough statistical power to detect an effect of time (despite using a similar number of participants and responses per participant as other serial recall experiments that supposedly provided evidence for time-based forgetting), the model-based analysis is less susceptible to such arguments based on power. That is, as Figure 1 (left column) shows, the time-based version of SIMPLE can predict virtually any extent of fanning of the serial position curves by adjusting the temporal attention weight. It follows that if time had even a slight effect, this should have been detectable by an improved fit of the model if temporal attention was greater than zero.

Finally, although our experiments necessarily involved a finite selection of retrieval times, they covered a considerable range (from 6 to 14 sec) and comfortably bracketed the retrieval times that are typical of unpaced serial recall of six-item lists (around 7 sec; e.g., Maybery, Parmentier, & Jones, 2002). Nonetheless, it is possible that time might be causally involved in forgetting outside the range of times examined in our experiments. However, there is little theoretical reason to expect time to have an effect beyond the longest retrieval times observed here: Time-based theories either expect the effect of time to be uniform across all delays (e.g., SIMPLE; Brown et al., 2002) or they expect time-based decay to have been largely completed within the present range of retrieval times (e.g., Baddeley et al., 1975, offer an estimate of around 2 sec for decay to take place). In any case, we note that the temporal distinctiveness model we have examined, and perhaps others as well, cannot offer a time-based explanation of the extended primacy observed in serial recall without at the same time predicting that recall speed variations of the range used here would produce a very large effect on the serial position curves.

Time and Memory: Theoretical Implications

We presented evidence against all of the many accounts, whether or not they are formally expressed as models, that attribute a significant role to the passage of time in forgetting from short-term memory. By implication, our results support an event-based approach, without, however, permitting adjudication between potential event-based alternatives.

At the outset, we explored two event-based mechanisms, output interference and an analogue of a primacy gradient, and found their predictions to be identical. It follows that because our experiments confirmed those predictions, they cannot also differentiate between the two alternatives. Such empirical differentiation is clearly possible—for example, we mentioned earlier that primacy gradients have difficulty explaining the extensive recency in probed recall—but is beyond the scope of the present article. For related reasons, we also chose not to fit our results with any particular event-based theory. It would be a trivial matter (and hence uninformative) to

take, say, SOB (serial order in a box; Farrell & Lewandowsky, 2002) and show that its predicted serial position curves (1) resemble those in our experiments and (2) do not differ with recall speed.

Time and Memory: The Wider Context

It is valuable to place the present results into the wider context of other recent findings that have forced a reevaluation of the role of time, in this instance during encoding rather than during retrieval. One central pillar of empirical support for temporal distinctiveness theories has been the finding that items that are temporally isolated at encoding are recalled better than items that are crowded together in time (e.g., Neath & Crowder, 1990, 1996). Specifically, Neath and Crowder showed that if the temporal spacing between items increased across serial positions, terminal list items were recalled better than their companions on a control list in which interitem spacing remained constant. Conversely, for lists on which the temporal spacing decreased with serial position, early list items had an advantage over those on the control list. Because total presentation time and retention interval were constant for all presentation schedules, and because rehearsal was supposedly prevented by the extremely brief durations involved (Neath & Crowder, 1996), these results have played a prominent role in justifying computational instantiations of temporal distinctiveness (e.g., Brown & Chater, 2001; Brown et al., 2002).

However, two recent experiments by Lewandowsky and Brown (in press) may force a reevaluation of the effects of temporal isolation at encoding. Lewandowsky and Brown randomly varied interitem intervals across serial positions and then used those intervals to predict performance. Specifically, using recall of an item in position i as a dependent measure, the interval between the item in position $i - 1$ and the item in position i (called the *pre* interval), and the interval between i and $i + 1$ (*post* interval) were entered as predictors into a hierarchical regression analysis. According to temporal distinctiveness theories such as SIMPLE, the more temporally isolated an item is (i.e., the larger the pre and post intervals), the better it should be recalled. However, Lewandowsky and Brown's results showed that only the post interval had an effect on recall performance. Furthermore, when participants engaged in AS, thus preventing rehearsal during encoding, there were no effects of interval at all (i.e., the effect of the post interval was abolished). The overall pattern of results is incompatible with a temporal distinctiveness view and instead supports an event-based approach, according to which people use the interval between items to rehearse (or otherwise consolidate) the preceding item unless such rehearsal is prevented by AS.

Thus, the results reported here represent another recent instance in which findings that have previously been taken to identify a causal role of time in memory do not withstand further empirical scrutiny. This raises the question about what steps are now open to those who continue to

favor the view that time plays a causal role in short-term memory. In our view, the onus is now on time-based theoreticians either (1) to show that a time-based theory can handle the present results (and those of Lewandowsky & Brown, in press) better than an event-based alternative or (2) to provide other, new empirical evidence that reintroduces the need for time as an explanatory concept. The former possibility seems somewhat unlikely, given that the purely temporal version of SIMPLE—which is an extremely powerful theory; see Brown et al. (2002)—was unable to accommodate our results (or those of Lewandowsky & Brown, in press; see their article for details). The latter possibility cannot be ruled out, although any new data on the role of time would, of course, require reconciliation with the present results.

REFERENCES

- BADDELEY, A. D. (1976). *The psychology of memory*. New York: Basic Books.
- BADDELEY, A. D. (1986). *Working memory*. Oxford: Oxford University Press, Clarendon Press.
- BADDELEY, A. D., CHINCOTTA, D., STAFFORD, L., & TURK, D. (2002). Is the word length effect in STM entirely attributable to output delay? Evidence from serial recognition. *Quarterly Journal of Experimental Psychology*, **55A**, 353-369.
- BADDELEY, A. D., & HITCH, G. (1993). The recency effect: Implicit learning with explicit retrieval? *Memory & Cognition*, **21**, 146-155.
- BADDELEY, A. D., THOMSON, N., & BUCHANAN, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning & Verbal Behavior*, **14**, 575-589.
- BEAMAN, C. P. (2002). Inverting the modality effect in serial recall. *Quarterly Journal of Experimental Psychology*, **55A**, 371-389.
- BEAMAN, C. P., & MORTON, J. (2000). The separate but related origins of the recency effect and the modality effect in free recall. *Cognition*, **77**, B59-B65.
- BJORK, R. A., & WHITTEN, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, **6**, 173-189.
- BOROWIAK, D. S. (1989). *Model discrimination for nonlinear regression models*. New York: Marcel Dekker.
- BROWN, G. D. A., & CHATER, N. (2001). The chronological organization of memory: Common psychological foundations for remembering and timing. In C. Hoerl & T. McCormack (Eds.), *Time and memory: Issues in philosophy and psychology* (pp. 77-110). Oxford: Oxford University Press.
- BROWN, G. D. A., & HULME, C. (1995). Modeling item length effects in memory span: No rehearsal needed? *Journal of Memory & Language*, **34**, 594-621.
- BROWN, G. D. A., NEATH, I., & CHATER, N. (2002). *SIMPLE: A local distinctiveness model of scale-invariant memory and perceptual identification*. Manuscript submitted for publication.
- BROWN, G. D. A., PREECE, T., & HULME, C. (2000). Oscillator-based memory for serial order. *Psychological Review*, **107**, 127-181.
- BROWN, G. D. A., VOUSDEN, J. I., MCCORMACK, T., & HULME, C. (1999). The development of memory for serial order: A temporal-contextual distinctiveness model. *International Journal of Psychology*, **34**, 389-402.
- 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.
- CAPLAN, D., ROCHON, E., & WATERS, G. S. (1992). Articulatory and phonological determinants of word length effects in span tasks. *Quarterly Journal of Experimental Psychology*, **45A**, 177-192.
- CAPLAN, D., & WATERS, G. S. (1994). Articulatory length and phonological similarity in span tasks: A reply to Baddeley and Andrade. *Quarterly Journal of Experimental Psychology*, **47**, 1055-1062.

- CONRAD, R., & HILLE, B. A. (1958). The decay theory of immediate memory and paced recall. *Canadian Journal of Psychology*, **12**, 1-6.
- COWAN, N., DAY, L., SAULTS, J. S., KELLER, T. A., JOHNSON, T., & FLORES, L. (1992). The role of verbal output time in the effects of word length on immediate memory. *Journal of Memory & Language*, **31**, 1-17.
- COWAN, N., NUGENT, L. D., & ELLIOTT, E. M. (2000). Memory-search and rehearsal processes and the word length effect in immediate recall: A synthesis in reply to Service. *Quarterly Journal of Experimental Psychology*, **53A**, 666-670.
- COWAN, N., NUGENT, L. D., ELLIOTT, E. M., & GEER, T. (2000). Is there a temporal basis of the word length effect? A response to Service (1998). *Quarterly Journal of Experimental Psychology*, **53A**, 647-660.
- COWAN, N., SAULTS, J. S., & BROWN, G. D. A. (2004). On the auditory modality superiority effect in serial recall: Separating input and output factors. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 639-644.
- COWAN, N., SAULTS, J. S., ELLIOTT, E. M., & MORENO, M. (2002). Deconfounding serial recall. *Journal of Memory & Language*, **46**, 153-177.
- COWAN, N., WOOD, N. L., NUGENT, L. D., & TREISMAN, M. (1997). There are two word-length effects in verbal short-term memory: Opposed effects of duration and complexity. *Psychological Science*, **8**, 290-295.
- CRAIK, F. I. M., GOVONI, R., NAVEH-BENJAMIN, M., & ANDERSON, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, **125**, 159-180.
- CROWDER, R. G. (1976). *Principles of learning and memory*. Hillsdale, NJ: Erlbaum.
- DOSHER, B. A. (1999). Item interference and time delays in working memory: Immediate serial recall. *International Journal of Psychology*, **34**, 276-284.
- DOSHER, B. A., & MA, J. J. (1998). Output loss or rehearsal loop? Output-time versus pronunciation-time limits in immediate recall for forgetting-matched materials. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 316-335.
- DUNCAN, M., & LEWANDOWSKY, S. (2003). *The time-course of response suppression for within- and between-list repetitions*. Unpublished manuscript.
- ELVEVÅG, B., WEINBERGER, D. R., & GOLDBERG, T. E. (2001). Short-term memory for serial order in schizophrenia: A detailed examination of error types. *Neuropsychology*, **15**, 128-135.
- FARRELL, S., & LEWANDOWSKY, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review*, **9**, 59-79.
- FRIEDMAN, W. J. (2001). Memory processes underlying humans' chronological sense of the past. In C. Hoerl & T. McCormack (Eds.), *Time and memory: Issues in philosophy and psychology* (pp. 139-167). Oxford: Oxford University Press.
- GLENBERG, A. M., & SWANSON, N. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 3-24.
- HENSON, R. N. A. (1998). Short-term memory for serial order: The start-end model. *Cognitive Psychology*, **36**, 73-137.
- HENSON, R. N. A. (1999). Positional information in short-term memory: Relative or absolute? *Memory & Cognition*, **27**, 915-927.
- LAMING, D. (1992). Analysis of short-term retention: Models for Brown-Peterson experiments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1342-1365.
- LEWANDOWSKY, S. (1999). Redintegration and response suppression in serial recall: A dynamic network model. *International Journal of Psychology*, **34**, 434-446.
- LEWANDOWSKY, S., & BROWN, G. D. A. (in press). Serial recall and presentation schedule: A micro-analysis of local distinctiveness. *Memory*.
- LEWANDOWSKY, S., & FARRELL, S. (2000). A redintegration account of the effects of speech rate, lexicality, and word frequency in immediate serial recall. *Psychological Research*, **63**, 163-173.
- LEWANDOWSKY, S., & LI, S.-C. (1994). Memory for serial order revisited. *Psychological Review*, **101**, 539-543.
- LEWANDOWSKY, S., & MURDOCK, B. B., JR. (1989). Memory for serial order. *Psychological Review*, **96**, 25-58.
- LI, S.-C., & LEWANDOWSKY, S. (1993). Intra-list distractors and recall direction: Constraints on models of memory for serial order. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 895-908.
- LI, S.-C., & LEWANDOWSKY, S. (1995). Forward and backward recall: Different retrieval processes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 837-847.
- LOVATT, P., AVONS, S. E., & MASTERSON, J. (2000). The word-length effect and disyllabic words. *Quarterly Journal of Experimental Psychology*, **53A**, 1-22.
- LOVATT, P., AVONS, S. E., & MASTERSON, J. (2002). Output decay in immediate serial recall: Speech time revisited. *Journal of Memory & Language*, **46**, 227-243.
- MAYBERY, M. T., PARMENTIER, F. B. R., & JONES, D. M. (2002). Grouping of list items reflected in the timing of recall: Implications for models of serial verbal memory. *Journal of Memory & Language*, **47**, 360-385.
- MAYLOR, E. A., VOUSDEN, J., & BROWN, G. D. A. (1999). Adult age differences in short-term memory for serial order: Data and a model. *Psychology & Aging*, **14**, 572-594.
- MCCORMACK, T., BROWN, G. D. A., VOUSDEN, J., & HENSON, R. N. A. (2000). Children's serial recall errors: Implications for theories of short-term memory development. *Journal of Experimental Child Psychology*, **76**, 222-252.
- MCGEOCH, J. A. (1932). Forgetting and the law of disuse. *Psychological Review*, **39**, 352-370.
- MURDOCK, B. B. (1995). Developing TODAM: Three models for serial-order information. *Memory & Cognition*, **23**, 631-645.
- NAIRNE, J. S. (1990). A feature model of immediate memory. *Memory & Cognition*, **18**, 251-269.
- NAIRNE, J. S. (2002). Remembering over the short term: The case against the standard model. *Annual Review of Psychology*, **53**, 53-81.
- NAIRNE, J. S., NEATH, I., & SERRA, M. (1997). Proactive interference plays a role in the word-length effect. *Psychonomic Bulletin & Review*, **4**, 541-545.
- NEATH, I. (1993). Contextual and distinctive processes and the serial position functions. *Journal of Memory & Language*, **32**, 820-840.
- NEATH, I. (1999). Modelling the disruptive effects of irrelevant speech on order information. *International Journal of Psychology*, **34**, 410-418.
- NEATH, I., & CROWDER, R. G. (1990). Schedules of presentation and distinctiveness in human memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 316-327.
- NEATH, I., & CROWDER, R. G. (1996). Distinctiveness and very short-term serial position effects. *Memory*, **4**, 225-242.
- NEATH, I., & NAIRNE, J. S. (1995). Word-length effects in immediate memory: Overwriting trace decay theory. *Psychonomic Bulletin & Review*, **2**, 429-441.
- NG, M. L. H., & MAYBERY, M. T. (2002). Grouping in short-term verbal memory: Is position coded temporally? *Quarterly Journal of Experimental Psychology*, **55A**, 391-424.
- NOSOFSKY, R. M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology*, **43**, 25-53.
- PAGE, M. P. A., & NORRIS, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review*, **105**, 761-781.
- SERVICE, E. (1998). The effect of word length on immediate serial recall depends on phonological complexity, not articulatory duration. *Quarterly Journal of Experimental Psychology*, **51A**, 283-304.
- SERVICE, E. (2000). Phonological complexity and word duration in immediate recall: Different paradigms answer different questions—A comment on Cowan, Nugent, Elliott, and Geer. *Quarterly Journal of Experimental Psychology*, **53A**, 661-665.
- TURVEY, M. T., BRICK, P., & OSBORN, J. (1970). Proactive interference in short-term memory as a function of prior-item retention interval. *Quarterly Journal of Experimental Psychology*, **22**, 142-147.

WAUGH, N. C., & NORMAN, D. A. (1965). Primary memory. *Psychological Review*, *72*, 89-104.

WINGFIELD, A., & BYRNES, D. L. (1972). Decay of information in short-term memory. *Science*, *176*, 690-692.

NOTES

1. Note that although we do not focus on decay as an explanatory principle, it is possible for similar effects to arise out of simple trace decay (e.g., Baddeley & Hitch, 1993; Laming, 1992). Specifically, the ratios that determine performance in SIMPLE can be mirrored by item activations if items decay according to a power function.

2. We thank Graham Hitch for bringing this article to our attention.

3. We note that this procedure assumes independence of errors and hence should be interpreted with some caution.

4. This can have the consequence that, for both model and data, column totals slightly exceed unity. This is not a significant problem in cases like the present one where there were very few intrusions produced by participants.

(Manuscript received April 14, 2003;
revision accepted for publication August 26, 2003.)