Psychological Review
2007, Vol. 114, No. 3, 539–576

0033-295X/07/$12.00 DOI: 10.1037/0033-295X.114.3.539

A Temporal Ratio Model of Memory

Gordon D. A. Brown
University of Warwick

Ian Neath
Memorial University of Newfoundland

Nick Chater
University College London

A model of memory retrieval is described. The model embodies 4 main claims: (a) *temporal memory*—traces of items are represented in memory partly in terms of their temporal distance from the present; (b) *scale-similarity*—similar mechanisms govern retrieval from memory over many different timescales; (c) *local distinctiveness*—performance on a range of memory tasks is determined by interference from near psychological neighbors; and (d) *interference-based forgetting*—all memory loss is due to interference and not trace decay. The model is applied to data on free recall and serial recall. The account emphasizes qualitative similarity in the retrieval principles involved in memory performance at all timescales, contrary to models that emphasize distinctions between short-term and long-term memory.

Keywords: time, memory, model, distinctiveness

It is characteristic of scientific laws to hold over a wide range of temporal, spatial, or physical scales. One would be surprised, for example, if the Hooke–Newton law of gravitation held for 1-mg objects but not 1-g or 1-kg objects, or for distances of centimeters but not distances of meters. Although laws may break down at extremes (e.g., for subatomic weights), the default search is for scientific principles that are universal in that they apply over as wide a range of temporal and spatial scales as possible (e.g., Barenblatt, 1996). In developing models of human memory, however, it is widely assumed that different principles apply over different (short and long) timescales. Here, in contrast, we explore the extent to which common retrieval principles might apply to human memory over both short and long timescales.

In the first section, we motivate the approach through a review of data suggesting that many memory phenomena, such as serial position effects, appear at least qualitatively and sometimes quantitatively similar over a wide range of timescales. We refer to this as temporal *scale similarity*, and several examples are reviewed throughout the article. True scale invariance is seldom seen in the human memory data, nor is it predicted by the model we develop. However, we argue that any comprehensive model will need to account for scale-similar effects as well as data generally taken to require the assumption of different memory retrieval processes operating over different timescales. Such an account may require a shift in emphasis toward models that incorporate retrieval processes that operate over both short and long timescales.

The next section follows Murdock (1960) in suggesting that serial position effects in memory and absolute identification reflect a common mechanism. Specifically, we suggest that retrieval of items from memory in terms of their location along a temporal dimension is a discrimination problem equivalent to the identification of stimuli in terms of their position along a dimension such as weight, line length, or loudness. Memory items that occupy crowded regions of a temporal continuum are hard to retrieve for the same reason that weights, line lengths, or loudnesses are hard to identify when they are similar to other items to be identified.

A temporal distinctiveness model is then introduced. The model assumes that (a) episodic memories in multidimensional psychological space are located along a dimension representing temporal distance from the point of retrieval, (b) the retrievability of an item is inversely proportional to its summed confusability with other items in memory, and (c) the confusability of items along a temporal dimension is given by the ratio of the temporal distances of those items at the time of recall. The emphasis on temporal ratios rather than absolute temporal durations gives the model its core properties. We term the model *SIMPLE* (for scale-independent memory, perception, and learning). The model bears a family resemblance to several hitherto disparate approaches, including Murdock's (1960) distinctiveness theory; ratio-rule models of memory retrieval (e.g., Baddeley, 1976; Bjork & Whitten, 1974; Crowder, 1976; Glenberg & Swanson, 1986); nontemporal exemplar models (Nosofsky, 1986, 1992); the feature model (e.g., Nairne, 1990); earlier temporal distinctiveness and oscillator-based models of memory (e.g., G. D. A. Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999; Neath, 1993a, 1993b); and, more

Gordon D. A. Brown, Department of Psychology, University of Warwick, Coventry, England; Ian Neath, Department of Psychology, Memorial University of Newfoundland, St. John's, Newfoundland, Canada; Nick Chater, Department of Psychology, University College London, London, England.

This research was supported by Grant 88/S15050 from the Biotechnology and Biological Sciences Research Council (United Kingdom), Grants R000239351 and RES 000231038 from the Economic and Social Research Council (United Kingdom), and Grant F/215/AY from the Leverhulme Trust. We thank John Anderson, Alice Healy, Ben Murdock, and numerous colleagues for many helpful comments and discussions.

Code for the simulations reported in this article is available at www.warwick.ac.uk/staff/G.D.A.Brown/simple/.

Correspondence concerning this article should be addressed to Gordon D. A. Brown, Department of Psychology, University of Warwick, Coventry CV4 7AL, England. E-mail: g.d.a.brown@warwick.ac.uk

generally, models that place time at the heart of memory and learning (e.g., Gallistel, 1990; Gallistel & Gibbon, 2002).

The majority of the article is devoted to applications of the model to a number of serial recall and free recall phenomena.

Scale-Similar Effects in Memory and Identification

The account that we develop focuses on effects that hold over different temporal scales. What reason is there to believe that such an account is either necessary or possible? In addition to the general scientific preference for explanations that hold over many scales, an emphasis on phenomena that remain invariant over time has been central to recent models of timing behavior, such as scalar expectancy theory (e.g., Gibbon, Church, & Meck, 1984), and to time-based approaches to animal learning (Gallistel, 1990; Gallistel & Gibbon, 2000). Indeed, Gallistel and Gibbon (2002) argued that in the case of animal learning it is essential to account for "the time-scale invariance of the acquisition process, which we take to be the single most important quantitative fact about conditioning discovered in a century of experimental work" (p. 165). Here we focus on empirical phenomena in human memory that show either qualitative or quantitative similarities over different scales, and such data are addressed by simulation below. To the extent that such effects are observed, we argue, there is a need to identify memory retrieval principles that apply over different timescales. To anticipate: We do not here address all data that have been taken to implicate qualitatively different memory mechanisms for different timescales. We do, however, suggest that there are striking continuities in memory over different timescales and that these scale similarities need to be addressed by models of memory (see Chater & Brown, 1999).

First, effects of serial position are often suggestive of scale-similar mechanisms. Bowed serial position curves, showing reduced memory for midseries items, are obtained in free recall, serial learning, probed serial recall, location memory, and some recognition memory tasks (for reviews, see Crowder, 1976; Lansdale, 1998; McGeoch & Irion, 1952; Murdock, 1974). Serial position effects are also found in retrieval from long-term memory (Baddeley & Hitch, 1977; Bjork & Whitten, 1974; Glenberg, Bradley, Kraus, & Renzaglia, 1983; Healy, Havas, & Parker, 2000; Healy & Parker, 2001; Nairne, 1991; Pinto & Baddeley, 1991; Roediger & Crowder, 1976; Watkins, Neath, & Sechler, 1989). Serial position effects in rather different tasks, such as order reconstruction, show strong similarities over very different timescales (e.g., for the dimensions of position within list and list within trial: Nairne, 1991; data below). Neath and Brown (2006) noted that when the position of items separated by 50 ms in a list must be remembered or the day of week on which an event occurred must be recalled (Huttenlocher, Hedges, & Prohaska, 1992), the observed serial position curves are qualitatively similar although timescales vary by six orders of magnitude. Similar effects are seen in data from grouping experiments, where items at the beginning and end of each group are better recalled, echoing the primacy and recency effects for the list as a whole (Frankish, 1985, 1989; Ryan, 1969a, 1969b; Hitch, Burgess, Towse, & Culpin, 1996). In recognition memory, serial position effects appear similar over a range of timescales when rehearsal is discouraged (Wright et al., 1990).

The most detailed evidence has come from the study of recency effects. Although recency effects disappear after a filled retention interval (Glanzer & Cunitz, 1966; Postman & Phillips, 1965), the effect reappears if the spacing between presented items is increased (Bjork & Whitten, 1974) and is seen when retrieval from long-term memory is required (Baddeley & Hitch, 1977; Pinto & Baddeley, 1991; Sehulster, 1989). More generally, the size of the recency effect appears to depend on the log of the ratio between the interpresentation interval between the items and the retention interval (Glenberg et al., 1980; Nairne, Neath, Serra, & Byun, 1997), at least to a remarkably close degree (see also Baddeley, 1976; Bjork & Whitten, 1974). Empirically, the ratio rule means that it would be impossible for an observer to judge from the size of a recency effect whether that recency effect arose from recall of a list of items presented 10 s apart and followed by a 50-s retention interval or a list of items presented 1 s apart and followed by a 5-s retention interval. Of course, different mechanisms could nonetheless underpin long-term and short-term recency effects (e.g., Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, & Usher, 2005; Davelaar, Haarmann, Goshen-Gottstein, & Usher, 2006); we address this debate below.

Second, power-law forgetting would be consistent with scale similarity in memory performance over time. If forgetting does follow a power law, then the probability of recall will depend on T^{-a}, where T is the amount of time since an episode was learned and a is a constant. To the extent that the time course of human memory loss does more or less closely follow a power law, as has been suggested by a number of researchers (see, e.g., Anderson & Schooler, 1991; Rubin & Wenzel, 1996; Wixted & Ebbesen, 1991, 1997), forgetting may be seen as scale independent.[1] Many researchers have, however, claimed that forgetting curves are not best described by a power law (for recent examples, see Chechile, 2006; Rubin, Hinton, & Wenzel, 1999; T. D. Wickens, 1999; cf. Myung, Kim, & Pitt, 2000); we explore the functional form of forgetting curves through simulation below.

Third, the proportion of errors produced for each serial position during serial learning remains constant even in the face of considerable variations in degree of learning, interpresentation interval, time between trials, familiarity or meaningfulness of the material to be remembered, or individual differences in the learners (H. W. Braun & Heymann, 1958; McCrary & Hunter, 1953); this is the *Hunter–McCrary Law*. Here the shape of the error distribution in serial learning provides the observer with little or no evidence about the absolute level of performance.

Many further data can be used to motivate the focus on scale-similar effects in memory. In serial recall or order reconstruction tasks, positional errors show orderly gradients such that items placed in an incorrect position are nonetheless likely to be recalled in approximately the correct location. This tendency, like serial position curves, is evident across timescales varying over many orders of magnitude, from milliseconds to weeks (see Estes, 1972; Huttenlocher et al., 1992; Nairne, 1991, 1992; Neath & Brown, 2006; see G. D. A. Brown et al., 2000, for a summary). Further evidence comes from the study of memory for relative recency.

[1] Power-law relationships are evidence for scale invariance in a way that other functions (such as logarithmic or exponential ones) are not, because a power-law relationship holds independently of the measurement scale. Thus, if y is proportional to x^a, the same relationship will hold even if all x values are multiplied by a constant (although the constant of proportionality will change, a will not).

Underwood (1977) asked participants to recall the dates of events that occurred between 4 months and 7.5 years in the past, and found that the greater the time separating two events was, the less likely those events were to be recalled in the wrong relative order. Hacker (1980) found similar effects for items separated by just a few seconds (see also Muter, 1979). There is also evidence for scale similarity over the time frame of recall. Maylor, Chater, and Brown (2001) asked participants to recall events from the past day, week, or year. The cumulative response probabilities were indistinguishable across the three conditions. Finally, scale similarity is also strikingly evident in recall-order effects in free recall (Howard & Kahana, 1999): The lag-recency effect, which is the tendency for items that were presented contiguously to be recalled contiguously (with an additional forward bias), is observed in both normal and continual distractor free recall.

In summary, many important properties of memory are temporally scale similar, in that qualitatively similar effects are evident at many different timescales (see also Melton, 1963; Nairne, 1992, 1996). Such phenomena appear to call for models of memory that can explain them. The attempt to construct such a model need not involve concluding that there are no differences in memory retrieval processes over time but rather implies a different starting point. In the study of human and animal timing, for example, the default assumption underpinning many models is of timescale invariance (scalar timing), which is seen as a core explanandum. Evidence for scale dependence in specific situations is seen as informative in the context of the default assumption of scale independence (see Buhusi & Meck, 2005, for a recent review). Here we explore a similar approach in the context of human memory, where the default assumption in the extant literature is, in contrast, one of complete temporal scale dependence (separate processes operating over short and long timescales).

Model Assumptions

The model we develop is underpinned by four key assumptions. First, it is assumed that memory traces can be seen as located and individuated at least partly in terms of their position along a temporal continuum receding from the present into the past. This time line is logarithmically compressed, such that recent locations are more easily discriminable from one another than are more temporally distant locations (cf. Bjork & Whitten, 1974; Crowder, 1976).

The second assumption is that memory retrieval can be viewed as a discrimination problem; a memory trace will be easy to retrieve to the extent that it is easily discriminable from its nearby neighbors in psychological space. Thus, memory retrieval is seen as akin to absolute identification (cf. Murdock, 1960), and serial position curves in memory and identification are assumed to have a common origin.

Third, the confusability and hence discriminability of memory traces along the temporal continuum is assumed to depend on the ratio of their temporal distances from the time of retrieval. Finally, it is assumed that the probability that any given memory can be retrieved will be inversely related to its summed confusability with other memories. We now describe each of these assumptions in more detail; the Appendix provides a worked example.

Temporal Memory

The model assumes that memories, at least of the type examined in conventional serial and free recall tasks, are separate episodic traces, individuated and located in terms of their position along a temporal dimension, such that recent items occupy more discriminable (less confusable) locations along the dimension and hence are more retrievable than distant items. The model does not suggest that the temporal dimension is the only one along which items are represented; rather, as in many exemplar models of categorization and identification as well as some models of short-term memory, it is assumed that items can be represented in terms of their location within a multidimensional space. Here we additionally assume that compressed temporal distance is one important dimension in such a space. The addition of a temporal distance dimension allows us to apply principles of the type previously explored in categorization and identification models (e.g., Nosofsky, 1986) to serial and free recall.

There is considerable empirical evidence that time is an important dimension underpinning memory organization and retrieval (e.g., Gallistel, 1990; see G. D. A. Brown & Chater, 2001, for a recent review and Friedman, 2001, for an alternative perspective). Much of this evidence is discussed below. The specific model we describe instantiates the analogy of a line of telephone poles developed by Crowder (1976) to illustrate the ratiolike mechanism proposed by Bjork and Whitten (1974):

> The items in a memory list, being presented at a constant rate, pass by with the same regularity as do telephone poles when one is on a moving train. The crucial assumption is that just as each telephone pole in the receding distance becomes less and less distinctive from its neighbors, likewise each item in the memory list becomes less distinctive from the other list items as the presentation episode recedes into the past. Therefore, retrieval probability is being assumed to depend on discriminability of traces from each other. (Crowder, 1976, p. 462; see also Baddeley, 1976; Glenberg et al., 1983; Hitch, Rejman, & Turner, 1980; Tan & Ward, 2000)

To make this concrete, Figure 1A shows the temporal schedule of presentation of a five-item list (A-B-C-D-E). Items are presented at a rate of one item per second, and recall is assumed to occur 2 s after presentation of the last item. Thus, the actual temporal distances of the items' memory representations at the time of recall range from 6 s (Item A) to 2 s (Item E). At encoding, items are assumed to be associated with the state of an internal temporal–contextual signal at the time the item is encountered (this is just how models such as oscillator-based associative recall [OSCAR; G. D. A. Brown et al., 2000] function).

Figure 1B shows how the positions of the memories will appear compressed from the temporal perspective of retrieval (specifically, after logarithmic transformation of the temporal distances; right justified to "the present"). In other words, $M_i = \log(T_i)$, where M_i is the location of item i on the psychological dimension and T_i is the temporal distance of the item at the point of recall. Note that the more temporally distant items now appear closer together and are hence less discriminable along the retrieval dimension. The exact determination of confusability is described in the next section.

At retrieval, performance is assumed to depend on the position of the target item's position along the temporal dimension (G. D. A. Brown et al., 2000). For example, suppose Item C is to

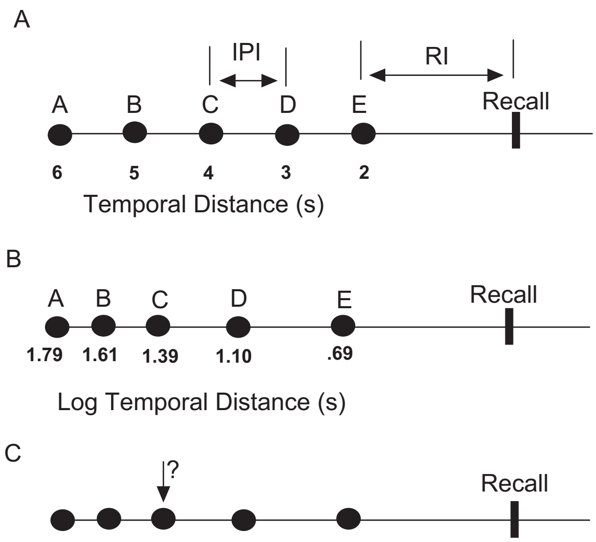

Figure 1. Illustration of the application of SIMPLE (scale-independent memory, perception, and learning model) to memory retrieval. IPI = interpresentation interval; RI = retention interval.

be retrieved. This item's location is illustrated by the arrow in Figure 1C. The probability of correct recall, given this location, will then be influenced by how discriminable the target item is from other items, that is, by the crowdedness of the local temporal neighborhood.

Note that Figure 1 (a) shows only the temporal dimension, and that other dimensions are also assumed to be important in accounting for nontemporal similarity effects in simulations below, and (b) represents only the simplified case where items are recalled after a fixed retention interval. The changing temporal perspective of recall (e.g., in serial order memory tasks) is considered below.

Memory, Discrimination, and Identification

Our second assumption is that memory retrieval is akin to discrimination and subsequent identification. This allows us to follow Murdock (1960) in linking identification to serial and free recall. The suggestion is that the same principles govern the discriminability (and hence retrievability) of items in memory as govern the discriminability of stimuli from one another in absolute identification paradigms (see also Miller, 1956, for discussion of capacity limitations in memory and identification). More specifically, serial position effects are assumed to arise for the same reasons in memory and in identification (Murdock, 1960), and hence the same underlying principles should be relevant to both absolute identification and free recall.

The key relevant data come from absolute identification tasks. In a typical absolute identification experiment, participants are exposed to a set of stimuli arrayed along some dimension such as pitch, weight, area, or amplitude (e.g., nine tones of different frequencies). A label is associated with each stimulus. The labels may be numbers (e.g., 1 through 9), with the number for each item corresponding to the item's ordinal position on the continuum, or may be arbitrary (e.g., the names of different colors may be associated with the different stimuli, in which case the task becomes more akin to paired-associate learning). Participants are then exposed to individual stimuli in random order and required to identify them with the correct label. Feedback regarding the correct response is normally given after each trial.[2]

The core of the absolute identification task is its requirement that items be identified from one another in virtue of their position along a dimension (weight, loudness, etc.). Items that occupy nearby locations along the dimension will be difficult to discriminate from one other, will tend to be confused with one another, and hence will be difficult to identify. Can memory retrieval be viewed in the same light? A number of well-established absolute identification effects mirror effects seen in memory retrieval, consistent with a similar discrimination process being involved.

First, as emphasized by Murdock (1960), bowed serial position curves are observed in absolute identification tasks as in many recall tasks (Murdock, 1960, focused on serial learning tasks with their characteristic large primacy and smaller recency effects). Bowed serial position curves are seen whether the relevant dimension is amplitude or weight (e.g., Murdock, 1960), frequency (e.g., Neath, Brown, McCormack, Chater, & Freeman, 2006; Stewart, Brown, & Chater, 2005), line length (e.g., Bower, 1971; Kent & Lamberts, 2005; Lacouture & Marley, 2004), area (Eriksen & Hake, 1957), position along a semantic continuum (DeSoto & Bosley, 1962; Pollio & Deitchman, 1964, cited in Bower, 1971), spatial position (Ebenholtz, 1963; Jensen, 1962), brightness (Bower, 1971), or numerosity (Neath et al., 2006). Of particular relevance to the present model, the same serial position curves are found when the stimuli to be identified are temporal durations (G. D. A. Brown, McCormack, Smith, & Stewart, 2005; Elvevåg, Brown, McCormack, Vousden, & Goldberg, 2004; Lacouture, Grondin, & Mori, 2001; McCormack, Brown, Maylor, Richardson, & Darby, 2002). The observed serial position curves are typically near-symmetrical when the stimuli to be identified are spaced approximately equidistantly on the psychological scale. Such symmetry contrasts with the asymmetrical serial position curves typically observed in free recall (large recency; small primacy). However, absolute identification tasks do show asymmetrical serial position curves, reminiscent of those seen in free recall, when the items to be identified are not evenly spread in psychological space. Panels A and B of Figure 2 show the serial position curves obtained in absolute identification when the stimuli to be identified (temporal durations, in the case illustrated) are either positively skewed (Figure 2A) or negatively skewed (Figure 2B). In both cases the shortest temporal duration was 100 ms and the longest was 900 ms; the distribution of durations within that range is illustrated in the insets of each panel (the data are a replotted subset of those reported in G. D. A. Brown et al., 2005). It is clear that absolute identification is less accurate in more crowded regions of stimulus space, with an additional advantage for end-series items (assumed to reflect their proximity to fewer items than midseries items). In particular, the positively skewed distribution

[2] Absolute identification is typically characterized by a near absence of performance improvement over large amounts of learning (e.g., Shiffrin & Nosofsky, 1994), and this, as well as the procedure involving use of unidimensional stimuli and an ordered response continuum, distinguishes it from paired-associate learning. Absolute identification and paired-associate learning are distinguished from absolute judgment, in which sensory magnitude judgments are given without feedback, although an associative component may be involved in absolute judgment (Haubensak, 1992; Wedell, 1996; but see Parducci, 1992).

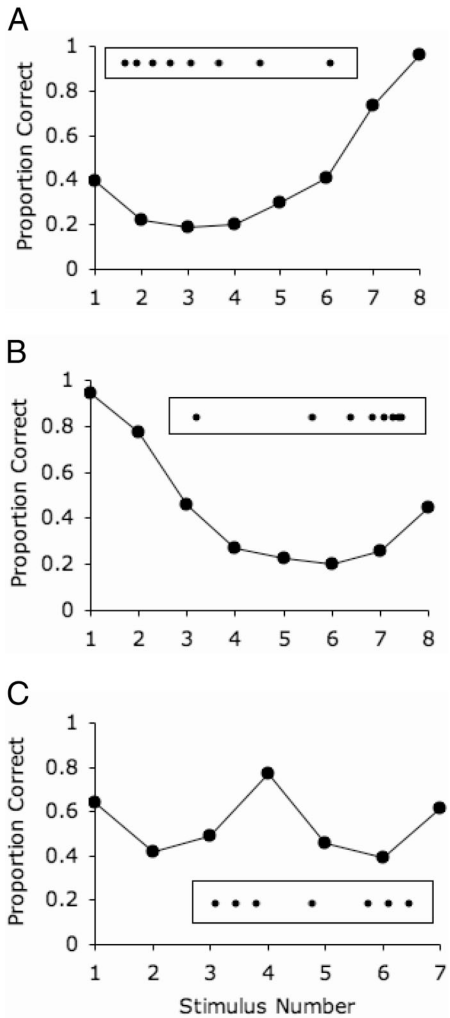

Figure 2. Serial position curves obtained in absolute identification experiments. A: Curve analogous to recency effect. B: Curve analogous to primacy effect. C: Isolation effect.

of temporal durations, which is similar to the temporal distances of memory items at retrieval time shown in Figure 1B, gives rise to an analogue of large recency coupled with small primacy effects. Similar results obtain when the underlying dimension is frequency instead of time (Neath & Brown, 2006). The results appear consistent with the suggestion that asymmetrical serial position effects in recall and identification paradigms reflect the same underlying mechanisms of discrimination.

A second parallel between memory retrieval in serial and free recall paradigms on the one hand and absolute identification on the other concerns scale similarity. Scale similarity in memory retrieval was noted above; here we note corresponding evidence for scale similarity effects in identification. The central (and counterintuitive) point is that, provided individual pairs of items are discriminable, identification performance may be almost unaffected when the spacing of items along the perceptual scale is increased by a constant factor (e.g., Alluisi & Sidorsky, 1958; Eriksen & Hake, 1955; Garner, 1962; Miller, 1956; Pollack, 1952; Shiffrin & Nosofsky, 1994). Scale similarity is also evident in the serial position effects obtained in absolute identification experiments, for such curves have essentially the same form no matter what the range of the relevant perceptual dimension that is used.

Third, absolute identification and conventional free recall both show *isolation effects*, also referred to here as *distinctiveness effects*, such that individual items that occupy isolated locations along the relevant dimension are advantaged. To illustrate, Figure 2C shows representative performance on an absolute identification task when the middle item of the stimuli to be identified is isolated along a frequency dimension (data adapted from Neath et al., 2006; inset shows stimulus distribution). Similar isolation effects are observed when the dimension is numerosity, weight, or the length of a physical rod (Neath et al., 2006). The key finding is that the isolated item is more accurately identified than when it is not isolated (because its immediate neighbors are not so close to it along the crucial dimension). According to the present perspective, these isolation effects in identification mirror temporal isolation effects in free recall, which occur when a list item that is either preceded or followed by a relatively long temporal gap at presentation is better recalled (G. D. A. Brown, Morin, & Lewandowsky, 2006). Temporal isolation effects in memory therefore appear consistent with the idea that memory retrieval in free recall is akin to discrimination and identification of items in terms of their position along a temporal distance continuum. The importance of near neighbors in determining item discriminability both allows us to develop a quantifiable notion of distinctiveness and distinguishes SIMPLE from previous distinctiveness models of identification and memory. For example, Murdock's (1960) influential model assumes that the distinctiveness of items is determined by their psychological distances from all items in the set of to-be-discriminated items (*global distinctiveness*), whereas SIMPLE (like previous exemplar-based models of absolute identification) proposes in contrast that distinctiveness is given primarily by an item's psychological distances only from other items that are nearby in psychological space (*local distinctiveness*). The local distinctiveness assumption underpins isolation effects in both absolute identification (Bower, 1971; Neath et al., 2006) and recall. Neath and Brown (2007) discuss the relationship between the approach described here and various other notions of distinctiveness.

The importance of confusability with near neighbors highlights a further similarity between absolute identification and memory retrieval (specifically, serial recall). In both cases, misidentifications predominantly reflect items with similar dimensional values to the target item. In the case of memory for serial order, probed recall, serial recall, and order reconstruction tasks all show a tendency for items that are not recalled in their correct positions to be recalled in nearby positions (see G. D. A. Brown et al., 2000, for a review). The same tendency is evident in absolute identification of both temporal and nontemporal stimuli (e.g., G. D. A. Brown et al., 2005; Stewart et al., 2005).

In summary, we emphasize the assumption of continuity between absolute identification and memory retrieval. In both cases, items are encoded in terms of their location along some continuous dimension (e.g., frequency or amplitude in absolute identification; time in memory). In both cases, serial position curves reflect an advantage for items near the ends of the continuum. In both cases, the retrievability of an item's trace depends on the confusability of that trace's position along the relevant continuum. Indeed, the model developed here to account for short-term memory and free recall can also be applied to isolation and serial position effects in absolute identification (Neath & Brown, 2006).

Confusability of Memories

The third core assumption of the model concerns item discriminability. The discussion above has assumed without argument that items occupying nearby positions along the temporal distance dimension will be less discriminable and hence less retrievable. In SIMPLE, the similarity of any two items' memory locations can be described in two (formally equivalent) ways. The first (exponential similarity distance metric) brings out the relation of SIMPLE to exemplar models of categorization and identification; the second expresses similarity as a ratio of temporal distances, thus allowing SIMPLE to be related to extant ratio models of memory. We describe these in turn.

In line with many other models, it is assumed that similarity $\eta_{i,j}$ between any two memory locations M_i and M_j falls off as a decreasing function of their separation in psychological space (Shepard, 1987). (Often in the simulations below, this will just be separation along the temporal distance dimension, but distance in multidimensional space is relevant in some simulations.) Thus, where memory traces differ along just a single dimension,

$$\eta_{i,j} = e^{-c|M_i - M_j|^{\alpha}}, \tag{1}$$

where c is a constant and α is 1.0 for an exponential function relating similarity to distance and 2.0 for a Gaussian function relating similarity to distance. Use of this function, which is widely used to relate similarity to distance in psychological space for separable stimuli (e.g., Nosofsky, 1986; Shepard, 1957, 1987), has the effect that items that are very close on a psychological scale have a similarity approaching 1.0 (because $\eta_{i,j} = 1$ when $|M_i - M_j|^{\alpha} = 0$, whereas items that have more psychologically distant representations from one another have a similarity that approaches zero as the psychological distance becomes greater (because $\eta_{i,j}$ tends to zero when the psychological distance $|M_i - M_j|^{\alpha}$ tends to infinity). The rate at which similarity reduces with psychological distance is given by the parameter c. The effect of using the exponential similarity–distance function is that the probability of correctly identifying a given item is determined most strongly by the similarity of its psychological scale value to those of its immediate neighbors (the local distinctiveness principle). More distant neighbors have less influence.

In applications to memory below, α was set to 1.0 (i.e., the function relating similarity to psychological distance was assumed to be exponential). This constrains and simplifies the model considerably and, in combination with the assumption that internal psychological magnitudes are logarithmically transformed temporal distances, allows the similarity of memory locations to be expressed very simply in terms of temporal distance ratios when no nontemporal dimension is involved.

Specifically, in the alternative formulation, the confusability of any two items based on their locations along the temporal dimension is just the ratio of their temporal distances raised to some power. The temporal distance of the more recent item is divided by the temporal distance of the less recent item. For example, the confusability of items that occurred 4 s and 5 s ago would be $(4/5)^c$, where c is the same free parameter as in the exponential similarity–distance formulation above. The confusability of items that occurred 1 s and 2 s ago would be smaller, being $(1/2)^c$. It is this emphasis on ratios of temporal durations in determining forgetting, rather than on absolute amounts of time, that gives the model its scale-similar properties and distinguishes it from trace decay models, at least as such models are traditionally conceived. The emphasis on ratios of temporal distances brings out the close relationship of SIMPLE to the temporal discrimination models of Baddeley (1976) and Bjork and Whitten (1974) in addition to other ratio models of memory (see Neath & Brown, 2007, for detailed discussion). The two formulations of similarity are formally identical (see Appendix).

Discriminability and Recall Probability

We assume, following a version of the similarity choice model (Luce, 1963; Shepard, 1957), that the discriminability D_i of a memory trace located at M_i will be inversely proportional to its summed similarity to other memory traces:

$$D_i = \frac{1}{\sum_{j=1}^{n}(\eta_{i,j})}, \tag{2}$$

where n is the number of items in the set of potentially retrievable items (often just the number of list items). This captures the idea that items will be more discriminable if they occupy relatively isolated locations on the temporal dimension (and indeed on other dimensions). In simple serial order memory paradigms where no omissions are possible, the probability $P(R_i)$ of recalling item i in its correct temporal position (given its location on the temporal dimension as a cue) will simply be its discriminability, hence,

$$P(R_i) = D_i = \frac{1}{\sum_{j=1}^{n}(\eta_{i,j})}, \tag{3}$$

or equivalently, using the ratio formulation of similarity, the probability $P(R_i)$ of recalling an item i in its correct position from a set of n items if the temporal distance of the item at the point of recall is T_i is given by

$$P(R_i) = \frac{1}{\sum_{j=1}^{n}\mathrm{Ratio}(T_i,T_j)^c}, \tag{4}$$

where Ratio(x,y) is the smaller of x and y over the larger and c is the free parameter described earlier.[3]

More generally, the probability of (incorrectly) recalling an item at M_j when the target memory has location M_i will be determined by the similarity of M_i to M_j relative to the similarity of M_i to all other values stored in memory. That is,

[3] From this point on we will use the exemplar model terminology, as this permits readier generalization to multiple dimensions, but it should be noted that alternative formulations in terms of temporal ratios are always possible where similarity only on the time dimension is concerned.

$$P(R_j|T_i) = \frac{(\eta_{i,j})^\gamma}{\sum_{k=1}^{n}(\eta_{i,k})^\gamma}, \tag{5}$$

where again n is the number of items in the set and $\eta_{i,j}$ is the similarity between M_i and M_j in memory. (Here the items are assumed to be equiprobable and response bias is ignored.) The parameter γ governs how deterministic responses are—high values of γ imply that participants will respond consistently in judging a given input; values near zero imply that participants will make highly variable responses with only a slight preference for the "correct response" (Ashby & Maddox, 1993). In the simulations below we set γ to 1.0, and the parameter can be ignored for present purposes.[4]

The model as described above can address probed serial recall and order reconstruction tasks, because in such tasks omissions cannot occur and items must be recalled in their correct serial positions. In modeling free recall, in contrast, additional mechanisms are needed to account for omission errors and the fact that items can be scored as correct even when they are recalled during the attempted recall of a different item. Omission errors are accommodated as follows. Intuition suggests that items that are difficult to discriminate should be most likely to be omitted. Any discriminabilities that fall below some threshold value will lead to omissions, and any that fall above the threshold will lead to overt recalls. Assuming some noise in activations or thresholds, this will have the overall effect of increasing recall probabilities that are already high and reducing recall probabilities for items whose recall probabilities are already low. We implement this via a sigmoid function such that if D_i is item discriminability calculated as above, recall probability is given by

$$P(R_i|D_i) = \frac{1}{1 + e^{-s(D_i - t)}}, \tag{6}$$

where t is the threshold and s (which gives the slope of the transforming function) can be interpreted as the noisiness of the threshold. For example, if t were set to .8 and s were very large, the transformation would approximate a system that recalled (with 100% probability) all items with relative strengths greater than .8 and omitted (with 100% probability) all items with strengths less than .8. As s becomes smaller, the transition from low to high recall probabilities becomes more gradual.[5] The possibility that in free recall an item may be correctly recalled in response to the attempted recall of another item is incorporated in the model as described below.

We now apply SIMPLE to a range of phenomena in serial and free recall from memory, with a particular focus on serial position effects.

Series 1: Serial Position Effects in Free Recall

In the first series of simulations, we apply the model to serial position data from a variety of paradigms. One question of primary theoretical interest is whether a unitary model can account for serial position effects of the type that have been taken as evidence for a separate short-term store. Further groups of simulations examine forgetting and proactive interference (Series 2) and serial recall (Series 3).

Serial Position Effects in Free Recall: Murdock (1962)

The serial position curve characteristic of immediate free recall includes large recency and smaller primacy effects at all list lengths. Representative data are reported by Murdock (1962), who presented lists of 10, 15, or 20 items at a rate of 2 s per item, and lists of 20, 30, or 40 items at a rate of 1 s per item, for free recall. These data are shown in Figure 3, along with the output of the model. In the case of free recall, we assume that the discriminability of each item is based on the item's time of occurrence relative to the time of retrieval. The crucial factor is the (log-transformed) value of the time that has elapsed between the learning of an item and the time of its retrieval.[6] As with most free recall tasks, however, the precise dynamics of participants' recall are not known, and therefore we made the simplifying assumption (discussed below) that the mean time of recall of an item was 15 s after the end of the list for the 10-item slower presented lists, 20 s and 25 s for the 15-item and 20-item slower presented lists, and 10 s, 15 s, and 20 s for the 20-, 30-, and 40-item faster presented lists. This figure could in principle be set independently of the observed data; here it was not possible to do so because the relevant time interval is not known.

As the task is free recall, we note that a list item may be recalled in response to an attempt to recall a different list item. For example, if the list to be recalled is A-B-C-D-E, a cue for the second item might retrieve B with probability .6, A with probability .1, C with probability .1, and so on. (This is analogous to the absolute identification case where B might be mistakenly identified as A or C on some percentage of occasions.) In free recall, an item will be scored correct whichever cue leads to its retrieval, and therefore in modeling we took the recall probability for a given item to be the sum of its recall probabilities over all retrieval attempts (subject to a maximum of 1).

In application of SIMPLE to free recall, there are three free parameters to estimate: c (temporal distinctiveness of memory representations), t (threshold), and s (threshold noise). Parameter estimates were those that minimized summed squared error (obtained using gradient descent methods). Here and throughout, parameters for simulations of empirical data are shown in Table 1.

The resulting fits (Figure 3) capture the key qualitative effects.[7] Both primacy and recency were obtained, although primacy was smaller in the model than in the data. The tendency for recency effects to be as observed but for primacy effects to be smaller in

[4] In the simple version of the model, the response determinism parameter will have effects redundant with those of c, and so there is no advantage in including the two parameters separately.

[5] The mechanism described has the consequence that zero or small values of discriminability can lead to higher predicted recall probabilities. In some applications, such as serial recall, this can occasionally result in a predicted sum of recall probabilities slightly greater than 1.0 for a given output position; rather than adopt a more complex thresholding mechanism we simply normalize or cap recall probabilities to 1.0 in such circumstances.

[6] Here and elsewhere we assume that the effective retention interval for a given item is measured from the offset, not the onset, of the item presentation.

[7] Here and elsewhere our focus is on capturing a range of qualitative effects across paradigms. We report R^2 values as a measure of fit despite the problems with the measure; direct log-likelihood calculations and model comparison are infeasible in most cases.

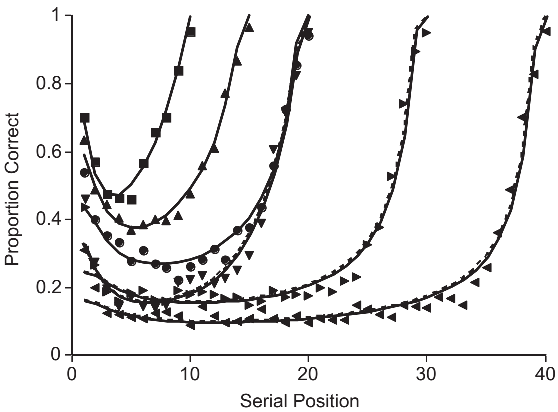

Figure 3. Serial position effects as a function of list length in single-trial free recall. Data adapted from Murdock (1962); solid lines show performance of SIMPLE with 10, 15, or 20 items presented at 2 s per item; dashed lines show performance of SIMPLE with 20, 30, or 40 items presented at 1 s per item. SIMPLE = scale-independent memory, perception, and learning model.

the model than in the data increases if shorter and arguably more plausible mean recall latencies (of 3, 4, and 5 s for the 10-, 15-, and 20-item 2-s lists, respectively; 6, 8, and 10 s for the 20-, 30-, and 40-item 1-s lists, respectively) are assumed (simulations not shown), consistent with the suggestion that the illustrated model fit is being driven by the need to fit primacy effects that in part reflect additional mechanisms, such as rehearsal, that are not incorporated into the model. We return to rehearsal and primacy effects below. The advantage for primacy items that is observed in the model arises because end items have fewer close neighbors than do middle-list items. This explanation in terms of "edge effects" has much in common with explanations proposed by Estes (e.g., 1972), Houghton (e.g., 1990), Treisman (1985), and others; the effects arise because items near the ends of the list have fewer near neighbors and are hence more discriminable.

Although this edge effect also contributes to superior performance on late-list items, an additional factor is required to explain the asymmetry in the curve. The large and extended recency arises because the temporal locations of items near the end of the list are more spread out and less confusable than are the codes for items earlier in the list (cf. the primacy and recency absolute identification conditions shown in Figure 2). This asymmetry follows directly from the logarithmic transformation carried out on the raw time values (cf. Figure 1B), because the transformation condenses large values (here, longer temporal distances) more than small values. We term this *Weberian compression* and discuss it in the context of the next simulation. In relation to such compression, we note that late-list items will benefit only if they are recalled early, because it is only under such circumstances that they benefit from their greater temporal distinctiveness. Such an assumption is consistent with the data, but we note that the model contains no mechanism to account for recall-order effects (e.g., Nilsson, Wright, & Murdock, 1975).

Abolition of Recency Effects After a Delay: Postman and Phillips (1965)

Early research emphasized the abolition of recency effects after a filled retention interval (Glanzer & Cunitz; 1966; Postman & Phillips, 1965); such evidence was initially seen as theoretically important in providing evidence consistent with a separate short-term store (although see Petrusic & Jamieson, 1978). Postman and Phillips (1965) had participants free recall a list of 20 items, presented at a rate of 1 per second, either immediately or after 30 s of interpolated activity. These data, replotted in Figure 4A, illustrate the classic large recency effect present for immediate recall but absent after a filled delay.

We examined the behavior of the model with the presentation schedule and retention intervals set as in the experiment. We assumed that the mean time of recall of an item was 20 s after the end of the experimentally imposed retention interval. There were the same three free parameters as in the previous simulation: c (temporal distinctiveness), t (threshold), and s (threshold noise). Parameter values were chosen to fit the immediate recall condition, and then the same parameter values were applied to the delayed recall condition so that resulting differences in the serial position function could be unambiguously attributed to changes in the temporal perspective of recall rather than changes in distinctiveness or threshold. Figure 4B shows the behavior of the model. Predicted recall after a delay (with the same parameter values) is evidently lower than observed; a better fit can be obtained if the threshold parameters are allowed to vary with retention interval. Most important, however, the model successfully predicts the abolition of recency with the filled delay, although as before, it does not fully capture the primacy evident in the data. As with the previous simulations, assumption of a shorter retention interval leads to less primacy in the model. Additionally, the basic model does not capture the preservation of primacy after a filled delay (Tan & Ward, 2000). We deal with these effects in turn.

Abolition of recency effects. Why does the abolition of recency after a delay occur in SIMPLE? The model instantiates a similar notion of discriminability to the one embodied in the telephone pole analogy. When recall is immediate, there is less Weberian compression of the scale on which the items are represented because only a short time has elapsed since the items' presentation. Thus, there is relatively little compression on that part of the scale where late-list items are represented, and substantial recency results. After a filled retention interval, in contrast, sufficient time has elapsed for the whole scale to have become compressed, and so end items lose their relative advantage almost completely. The model also predicts the observed reemergence of recency when list presentation is slow relative to the retention interval. Bjork and Whitten (1974) found that a recency effect could be observed even after a 30-s filled retention interval if sufficient time intervened between the presentation of each list item, and suggested that

> the necessary conditions for recency-sensitive retrieval from long-term memory can be specified by an empirical law of sorts based on the ratio of the temporal separation of successive to-be-remembered items (or sets of items) to the temporal delay from those items to the point of recall. (Bjork & Whitten, 1974, p. 189)

This led to the formulation of a Weber-like function relating the amount of recency to the log of the ratio of the duration of the interpresentation interval to the retention interval (see also Baddeley, 1976; Glenberg et al., 1983; Nairne et al., 1997). Thus, if the retention interval increases and interpresentation interval is held constant, as in the Postman and Phillips (1965) study, the ratio becomes smaller and recency decreases. Similar behavior occurs in SIMPLE: The relative advantage of recency items decreases grad-

Table 1
Parameter and R^2 Values for Simulations Reported in Text

Task	Figure	c	t	s	w_T	R^2
Immediate free recall						
10 items at 2 s per item	3	12.10	.53	8.98		.97
15 items at 2 s per item	3	10.82	.42	12.09		.98
20 items at 2 s per item	3	9.69	.35	16.19		.96
20 items at 1 s per item	3	10.45	.40	16.01		.96
30 items at 1 s per item	3	9.41	.30	21.12		.95
40 items at 1 s per item	3	12.03	.30	22.83		.96
Immediate and delayed free recall	4B	15.10	.41	14.06		.86
Recency of rehearsal	4C	19.97	.52	11.60		.99
Tone comparison						
Response time proportional	6	1.69				.65
Response time variable	6	1.52				.92
Proactive interference						
Model 1	7	1.5	.28	6.35		.90
Model 2	7	811	.39	26.50	.998	.84
Peterson forgetting	9	0.45	.18	23.30		.92
Forgetting and proactive interference						
One-parameter model	10B	1.25				.94
Three-parameter model	10C	0.48	.28	6.03		.95
Evidence against trace decay	11	1.92				.96
Release from PI						
By category shift	13A	1.5			.50	
By time	13B	4.0				
Power-law forgetting						
Fast	14	2	.75	5	.90	.98
Medium	14	4	.75	5	.90	.99
Slow	14	6	.75	5	.90	.97
Order reconstruction						
After 30 s	17	16.54				.96
After 4 hr	17	3.5×10^3				.91
After 24 hr	17	10.4×10^3				.68
List and position reconstruction	18	8.54			.91	.86
Serial recall						
Dissimilar items	20	8.90	.49	8.12	.95	.98
Similar items	20	8.90	.43	8.12	.95	.98
Isolation effects	21B	3.25			.12	.93
Serial recall						
Ungrouped	22	12.4	.69	5.8	.94	.89
Grouped	22	12.4	.69	5.8	.78	.89

Note. Parameter c signifies rate at which similarity between memory representations declines with distance, t signifies threshold, s signifies threshold noise, and w_T signifies attentional weight given to the global temporal dimension. PI = proactive interference.

ually as retention interval increases, because the spacing (and hence retrievability) of recency items relative to other list items reduces through Weberian compression as retention interval increases.

SIMPLE differs from simple ratio models in the introduction of the c parameter (so that the confusability of any two items in memory is not simply proportional to the ratio of the items' temporal distances but is instead the ratio raised to a power) and in its emphasis on the need to discriminate memory items from several near neighbors rather than just the closest. The latter property enables SIMPLE, unlike previous ratio models (Crowder, 1976), to address primacy effects.

Primacy effects. In simulating the Murdock (1962) and Postman and Phillips (1965) data, we observed less primacy in the model than in the data. Although there is some extended primacy in the model due to edge effects, there was no significant advantage for the first few items in the model. Must primacy effects then be due partly to encoding processes, especially for longer lists? Perhaps early-list items are more strongly encoded, and the SIMPLE model produces rather too little primacy because it does not incorporate such a factor? Here we argue for an alternative explanation: Primacy effects in free recall partly reflect rehearsal (e.g., Murdock & Metcalfe, 1978; Tan & Ward, 2000), and indeed Laming (2006) has shown that free recall sequences can be predicted (as probability distributions) from sequences of overt rehearsals. Thus, an important limitation of the model presented here is that it does not accommodate rehearsal processes, and this limitation, rather than the absence of encoding differences for early-presented items, may account for the model's underprediction of primacy when rehearsal is permitted. Crowder (1976) summarizes evidence for the contribution of active encoding processes to primacy effects. However, SIMPLE makes a different prediction from the active encoding account of primacy effects. Active encoding accounts must predict that primacy effects will be reduced or abolished by rehearsal-preventing activity such as continual distraction for any list length, whereas SIMPLE predicts that

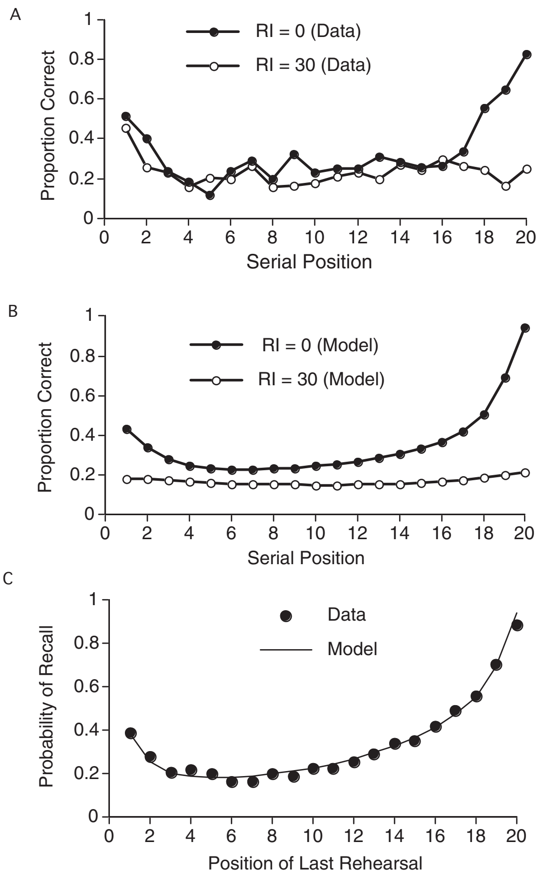

Figure 4. Immediate (retention interval [RI] = 0) or delayed (RI = 30) free recall. A: Data (adapted from Postman & Phillips, 1965). B: Behavior of SIMPLE (scale-independent memory, perception, and learning model). C: Observed probability of recall as a function of position of last rehearsal. Data adapted from Rundus (1971); solid line shows performance of SIMPLE.

primacy effects will be larger for short than for long lists and will survive even when active encoding is prevented. Consistent with this prediction, constraining participants to rehearse only the current item reduces but does not abolish primacy (Glanzer & Meinzer, 1967; Modigliani & Hedges, 1987; Tan & Ward, 2000). More generally, primacy effects still occur when rehearsal is limited by the nature of the task (Tzeng, 1973; Watkins et al., 1989; Wixted & McDowell, 1989; Wright et al., 1990) and indeed in animals without language (e.g., Harper, McLean, & Dalrymple-Alford, 1993). Rundus (1971) interpreted the evidence as consistent with the idea that other factors such as the "distinctiveness of the initial study items" (p. 65) are relevant. However, SIMPLE (in contrast to both active encoding and previous ratio-rule models) predicts that one source of primacy effects is a pure edge effect. This edge effect will, according to the model, be reduced for longer list lengths,[8] be evident under incidental learning conditions, and remain when active maintenance rehearsal is prevented.

Although SIMPLE does not incorporate rehearsal, it should be able to account for primacy effects that remain when the changes in temporal distinctiveness that are induced by rehearsal are removed or controlled (Brodie & Prytulak, 1975; G. D. A. Brown et al., 2000; Murdock & Metcalfe, 1978; Rundus, 1971; Tan & Ward, 2000). Even random rehearsal will tend to "telescope" the effective temporal distances of items at the point of recall at the end of the list, such that early-list items have shorter effective retention intervals than those given by the temporal distances of the items' initial presentations; there is ample evidence that controlled rehearsal will have this effect (e.g., Ornstein, Naus, & Stone, 1977; Rundus, 1971; Tan & Ward, 2000). The relative temporal distinctiveness of primacy items in free recall may also be increased owing to their relatively early recall (Murdock, 1974). In addition, the multiple and temporally distributed traces laid down by rehearsal (Modigliani & Hedges, 1987) may lead to an advantage for primacy items, for which there is opportunity for greater distribution of rehearsal, and a concomitant disadvantage for late-presented items, especially after a delay ("negative recency"), due to the massed nature (or different nature: Watkins & Watkins, 1974) of rehearsals of those items (Tan & Ward, 2000). These various factors, all of which may influence the temporal distinctiveness of items at the time they are retrieved, have not yet been completely disentangled in the empirical literature, but several involve rehearsal. We therefore examined the extent to which primacy is predicted by SIMPLE when just recency of rehearsal is taken into account.

If time since last rehearsal is the most important factor, as the most straightforward interpretation of SIMPLE suggests, the model should be able to predict the full extent of primacy observed when free rehearsal is allowed but the probability of item recall is described as a function of last rehearsal set. Rundus (1971) presented such data, as did Tan and Ward (2000). We averaged the data from Rundus (1971; Experiment 1) and the four relevant conditions of Tan and Ward (2000; Experiments 1 and 2).[9] The averaged data are reproduced in Figure 4C together with the behavior of SIMPLE.[10] In obtaining this fit, an average retention interval of 16 s was assumed. SIMPLE now produces the correct amount of primacy, consistent with the model's prediction that the extent of primacy will largely be determined by the temporal distance of the last rehearsal of an item rather than the temporal distance of that item's presentation. (Separate simulations found that the data sets of Rundus and of Tan & Ward, combined above, could be captured individually.)

Although SIMPLE predicts the extent of primacy reasonably well when the temporal distance of last rehearsal is taken into account, the simulated experimental conditions still only approximate reality because interference from traces other than most recent rehearsals is ignored and differential output time effects are not taken into consideration. In free recall, it is frequently observed that the most recent items are recalled first (an adaptive strategy according to SIMPLE, for this enables them to take advantage of Weberian compression), followed by primacy items particularly if recall is immediate, followed by other items. Precise protocol of output is somewhat variable, although a strong forward bias is

[8] This effect is also observed experimentally (see Atkinson & Shiffrin, 1971; Murdock, 1962; Postman & Phillips, 1965).

[9] Tan and Ward found small or absent effects of rate of presentation, or word frequency, when recall probability was considered as a function of last rehearsal set.

[10] For the Tan and Ward (2000) data, the averaging process involved weighting the results for the differing number of data points contributing to each position of last rehearsal.

normally evident (e.g., Laming, 1999; Murdock, 1974). Further simulations can be conducted to estimate the effective retention intervals (defined as time elapsed between last rehearsal of an item and recall of that item) for each item that would have given rise to the observed recall probabilities. Such models, not reported in detail here, ignore the possibility of output interference but nevertheless produce a good fit to the data. This result is unsurprising, as a separate parameter is available for each item, but lends strength to the claim that the SIMPLE model can give an improved account of primacy effects when the true effective retention intervals for each item are taken into consideration. However, the account remains limited in its implicit assumption that traces other than the most recently rehearsed trace for a given item do not cause interference; absent an explicit mechanism for rehearsal, the model is best interpreted as an account of memory under experimental conditions that preclude rehearsal.

In summary, SIMPLE predicts that in free recall some primacy will occur owing to the increased local distinctiveness of edge items even in the absence of rehearsal. This marks a clear difference from earlier ratio models. When rehearsal is possible, primacy effects increase owing to the relative reduction in effective retention intervals for primacy items. Other factors may increase the temporal distinctiveness of early compared with midlist items; such factors may include output order, the distributed rather than massed nature of the multiple traces laid down by rehearsal, or reduced encoding strength for late-list items. However, an important limitation of SIMPLE (albeit one shared by most other models, but cf. Laming, 2006) is its lack of a rehearsal-generation mechanism. We return to this issue in the General Discussion.

Apparent Departures From the Ratio Rule

The success of ratio-rule models has been seen as theoretically important because, to the extent that recall level is as predicted by a ratio rule, there is no need to postulate forgetting due to trace decay. However, a challenge comes from recent claims that the absolute amount of time since item learning, not just ratios of temporal intervals, influences recall probability (Cowan, Saults, & Nugent, 1997; see also Nairne et al., 1997). This challenge is therefore also of considerable significance not only to the SIMPLE temporal discrimination model but to any model that claims that the passage of absolute (rather than relative) time causes forgetting. Here we use the model to illustrate the general point that claims of absolute time effects are difficult to uphold unless uncontrolled effects of proactive interference (PI) can be excluded.

More specifically, we used the model to explore the data of Cowan et al. (1997), who suggested that the absolute amount of time that has passed before a memory test takes place can affect auditory memory performance even when the temporal schedule of item presentations and recall is held constant in ratio terms. Such an effect would violate scale similarity and suggest a role for absolute, rather than just relative, time in forgetting. To anticipate: The SIMPLE model is used to illustrate how the Cowan et al. findings can be understood in terms of PI rather than trace decay. In intuitive terms, this is because the relative importance of PI from earlier items increases as the interpresentation interval and retention interval become larger, even if these intervals remain in proportion to one another.

Cowan et al. (1997) presented participants with pairs of tones, the task being to say which tone was higher in frequency. The absolute time between the tones (retention interval) was varied. Previous studies found that performance decreased as retention interval increased. However, Cowan et al. included a novel manipulation of the time between the first tone in the pair and the last tone of the preceding pair of tones. This scheme is illustrated in Figure 5. Each solid vertical line represents the presentation of a single tone. In Case A, a 4-s between-trial gap separates the second tone of trial $n - 1$ from the first tone on trial n, and a 2-s retention interval separates the first tone of trial n from the second tone of trial n. This gives a ratio of 2:1 between the between-trial gap and the retention interval for trial n. In Case B, in contrast, each duration is 50% longer (6 s for the trial gap; 3 s for the retention interval), but the ratio remains the same: 2:1. It is therefore possible to examine the effects of varying the absolute amount of time both within and between pairs while holding the intrapair-to-interpair temporal ratio constant.

Cowan et al. (1997) did just this and found that performance reduces as a function of the absolute retention interval even when the ratio is held constant. Their results are illustrated in Figure 6, where we show just the case where the ratio between trial gap and retention interval was held constant at 2:1. Cowan et al. interpreted the effect of absolute time as evidence against the suggestion that the sole determinant of performance was the ratio of the between- and within-pair gaps, and concluded that trace decay also played a role in forgetting (although see Cowan, Saults, & Nugent, 2001, for considerations similar to those adduced here).

However, there are two ways in which a ratiolike model such as SIMPLE, in which there is no trace decay, could accommodate such a finding. First, an absolute amount of time is involved in the time between presentation of the second tone and the making of a response. The actual memory retrieval and comparison process must necessarily take place some finite amount of time after presentation of the second tone of the pair to be compared within a trial. This is illustrated in Figure 5, where the dashed vertical line reflects the actual time of retrieval of the first tone of a pair from memory. There is a short "response time," which is unlikely to increase in direct proportion with the retention interval. Second, Cowan et al. (1997) explicitly assumed that the influence of previous tones on previous trials would be negligible. In Figure 5, we have included the first tone of trial $n - 1$, which will, on average, be a constant amount of time prior to the second tone of that trial. Intuitively it is evident that the relative importance of PI from these earlier trials will become greater with the passage of absolute time; indeed we use this effect below to account for forgetting in Brown–Peterson tasks.

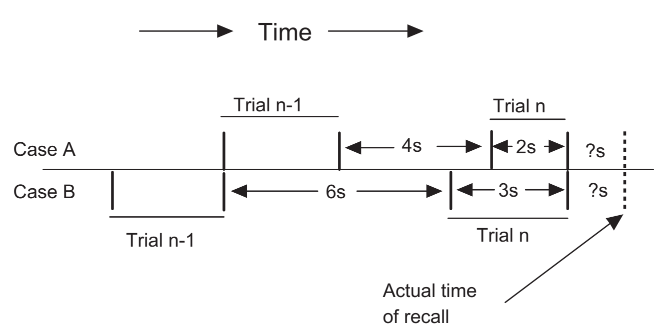

Figure 5. Diagrammatic illustration of the experimental procedure used by Cowan et al. (1997).

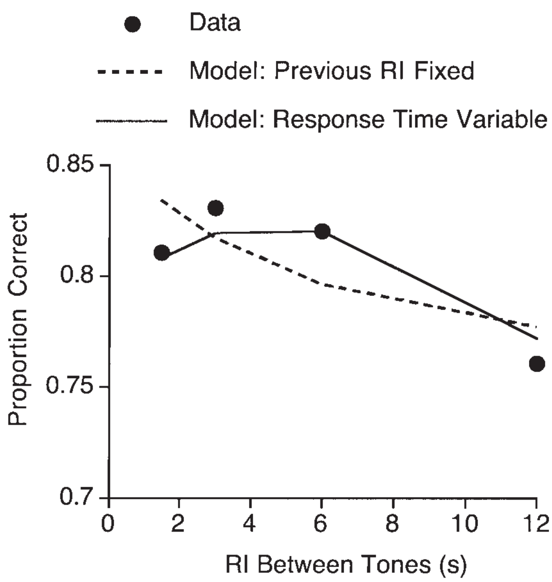

Figure 6. Proportion correct performance in a tone comparison task as a function of time between tones (data adapted from Cowan et al., 1997). Filled circles show the data; dashed line illustrates performance of SIMPLE (scale-independent memory, perception, and learning model) when response time is assumed fixed; solid line illustrates performance of SIMPLE when response time is free. RI = retention interval.

To confirm these intuitions, we used the SIMPLE model to explore the case illustrated in Figure 5, focusing on the importance of retention interval. We examined recall of the crucial first tone in the context of just one previous pair; the retention intervals were set to 1.5, 3, 6, and 12 s, as gave rise to the experimental data in Figure 6, and the corresponding between-trial gaps were set to 3, 6, 12, and 24 s, thus preserving a 2:1 ratio. We set the retention interval for the previous (nontarget) pair of items at a constant 10 s for the shortest retention interval and increased response time (defined here as the time between presentation of the second tone of the target pair and the time of memory retrieval and decision making) in proportion to the retention interval for the target pair (0.2 s as the smallest value). This ensured that any effect of absolute time would be due solely to the retention interval for the pair of tones preceding the target pair. Only the c parameter was free to vary; the fit shown in Figure 6 as a dashed line was obtained. A clear effect of absolute retention interval, of similar magnitude to that seen in the data, is observed. Thus, the apparent effect of absolute time may be due to PI from earlier trials.

In a further simulation, we allowed response time to be determined independently for each retention interval, and all other times were scaled in proportion to retention interval. In the Cowan et al. (1997) experiment, participants were given a maximum of 2 s to make a response, and so estimated response times were constrained to lie between 0.1 s and 2 s. A good fit was possible (see Figure 6), unsurprisingly given the number of parameters (five) in relation to the number of data points (four). More important, the response time estimates were in a psychologically plausible order, being 0.1 s, 0.1 s, 0.18 s, and 2.0 s for retention intervals of 1.5 through 12, respectively. Despite the overparameterization and the fact that the observed solution may be nonunique, the good fit indicates that incorporation of response time can allow an account of the data by a model with no trace decay. It is therefore possible that no temporal trace decay or other absolute time mechanism is needed to account for the results of Cowan et al.; we emphasize the more general point that investigations into the role of absolute time in forgetting must take account of all sources of proactive interference.

Discussion

The SIMPLE local distinctiveness model appears to give a reasonable qualitative account of several key findings concerning recency effects, their abolition with filled delay, and primacy effects in free recall. It makes different predictions from other ratiolike models, in that it predicts that primacy effects in free recall can arise at retrieval. In the simulations presented so far, forgetting of a single list is entirely due to interference; there is no time-based decay of traces. Some of the evidence that has been interpreted in terms of time-based forgetting does not falsify a temporal-discrimination interference-based account of the type presented here. Thus, SIMPLE illustrates how forgetting can occur owing to the passage of time alone even though no trace decay occurs (see also Nairne, 1996).

Series 2: Interference-Based Forgetting and Its Time Course

The previous simulations applied SIMPLE to serial position effects in free recall but skirted the issue of forgetting due to previous lists. The simulations in the present section emphasize the role of PI and its interaction with temporal factors. Many recent models fail to acknowledge the fact that previous trials in an experiment can greatly affect performance on a subsequent trial. Henson (1996) found that over 40% of intrusion errors in serial recall came from the list that immediately preceded the list to be recalled (see also Estes, 1991). Furthermore, many researchers have claimed that in the absence of PI, little or no forgetting will occur (Keppel & Underwood, 1962; Turvey, Brick, & Osborn, 1970; Underwood, 1957). Here we use SIMPLE to offer a perspective on PI effects.

Forgetting as a Function of Previous Lists: Underwood (1957)

Underwood (1957) collated the results of over a dozen published experiments, all of which examined memory performance after a 24-hr retention interval. These studies varied primarily in the number of lists that had been learned prior to the target list. The results, which are shown in Figure 7, were taken as clear evidence that "the greater the number of previous lists the greater the proactive interference" (Underwood, 1957, p. 53). When there is little or no PI, little forgetting occurs, even after 24 hr. This apparent lack of forgetting when there is no PI poses a major problem for many current models of memory.

SIMPLE can be used to address the basic data. We illustrate first with a "naive" model that treats each list as a single unit in memory. We then develop a more sophisticated account in which the amount of PI for an item depends on that item's distinctiveness in a two-dimensional space, where the two dimensions represent item-within-list position and list-within-trial position.

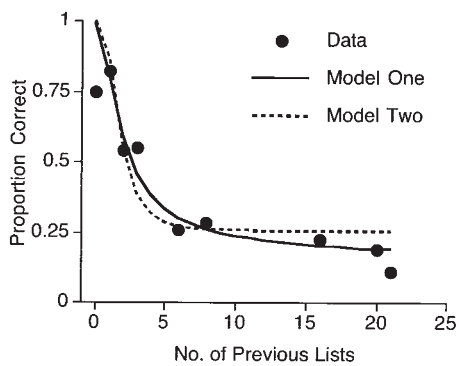

Figure 7. Proportion correct performance as a function of number of previous lists. Data adapted from Underwood (1957); lines show performance of two versions of SIMPLE (scale-independent memory, perception, and learning model).

The naive account illustrates the basic principles governing PI in the model. SIMPLE predicts that PI will be a major determinant of memory performance because any temporal retrieval cue for a list will be less effective when the number of lists occupying similar temporal positions to the target list increases. As a result of Weberian compression, previous lists may be quite strong candidates for retrieval on the basis of a temporal cue for the most recent (target) list. A current list is the most temporally distinctive because it has suffered the least amount of Weberian compression. However, any temporal retrieval cue intended to retrieve that list will also act as a partial cue for the previous list and (to a lesser extent) for the list before that. Thus, the major factor limiting performance will be the number of previous lists and their temporal separation from the current list. The account of PI is essentially similar to the explanation of recency effects; the same basic discrimination mechanisms are invoked in both cases.

To illustrate, we attempted to capture the basic features of the Underwood (1957) study. We assumed a 24-hr retention interval and a 60-s gap between lists. Each list was assumed to behave in the same way as single items in the studies described above (this amounts to an assumption that the main source of PI lies in discriminating lists, rather than items within a list, from each other). The results are shown in Figure 7 (solid line, labeled *Model One*).

As can be seen, performance is at almost 100% for a single list (because there is no PI). Performance then drops off sharply as the number of previous lists increases, because each list becomes less distinctive because of the increased number of competing lists. Addition of more and more proactively interfering lists has less and less additional effect, and PI has reached close to its maximum level after about three or four previous lists are taken into account. This is because of the sensitivity of the model to local neighborhood—neighbors have progressively less impact on discriminability as they become more temporally distant from the to-be-discriminated item, and so neighbors more than three or four items away are sufficiently temporally distant from, and hence dissimilar to, the target list that they exert negligible influence.

Although this basic model illustrates the mechanism of PI, it is doing little better than guessing. In any case, a more sophisticated account is needed to account for PI in the case where each list contains several separate items. A complete explanation must account for the separate effects of intraserial interference (the difficulty of discriminating a target item from other list items on the basis of a temporal retrieval cue) and interlist interference (where forgetting of a single list is faster when that list is preceded by other, proactively interfering lists). In particular, one important finding concerns the strong interference, during retrieval of a given item from the target list, from items that were presented, rehearsed, or recalled in the same within-list position on the previous trial during serial recall (e.g., Conrad, 1960; Estes, 1991). Can SIMPLE account for these data?

Several models of memory invoke hierarchical representations, in which items' positions along different dimensions, typically positional rather than temporal (e.g., item within list; list within trial), are represented simultaneously and independently (e.g., G. D. A. Brown et al., 2000; Burgess & Hitch, 1992, 1999; Henson, 1998b; Lee & Estes, 1981; Nairne, 1991). Within the present framework, we can conceive of items as being retrieved on the basis of their position in a two-dimensional space, where one dimension represents the within-list position of that item and the other dimension represents the temporal position of that item within a whole set of lists. This is illustrated in Figure 8, for the case of three lists of four items. In Figure 8, the vertical axis represents within-list position (1 through 4) and the horizontal axis represents the (log-transformed) amount of time between item presentation and a retrieval episode taking place immediately following the third and final list (shorter distances on the right). In SIMPLE, the retrievability of a given item will depend on its distance from its near neighbors in this two-dimensional space, in the same way as local neighborhood in a one-dimensional space (location along a simple temporal dimension) has governed performance in the simulations presented above. We assume that the distance of an item from any other item is simply the sum of its distances from that item along each of the two dimensions (i.e., we used a city block rather than euclidean metric). The probability of failing to discriminate items on the basis of their locations in two-dimensional memory space will be an exponential function of the distance between them, as before. It is evident that the discriminability of an item will depend on the closeness of both other list items and items that occupied similar positions within other lists. The balance between these will depend on the gap between lists relative to the interitem spacing within lists.

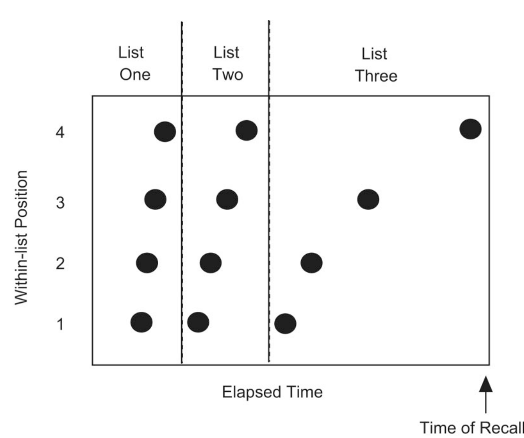

Figure 8. Illustration of the representation of 3 four-item lists in two-dimensional psychological space.

We examined PI in this two-dimensional model, again following the basic paradigm explored by Underwood (1957), although detailed data-fitting was not possible or appropriate given the variety of studies subsumed in the Underwood analysis. To explore the model's performance we assumed the retention interval was 24 hr, the 12-item lists were separated by 60 s (both of these assumptions being unchanged from the model above), and 1 s separated each within-list item.

A fourth parameter is required in this version of the model, to accommodate the possibility of selective attention being paid to particular dimensions of psychological space during recall. Attentional parameters are widely used in models of categorization (see, e.g., Nosofsky, 1992) and in intuitive terms can be thought of as stretching out the relevant psychological space in one dimension while simultaneously squashing it along another. Participants' performance in categorization tasks is well accounted for with the assumption that they can learn to pay selective attention to the stimulus dimension that is most task relevant. We introduced a new parameter, w_T, which specifies the attentional weight given to the global temporal dimension. The attention paid to the second dimension, here within-list position, is denoted w_P and is set to $(1 - w_T)$ to capture the notion that increasing attention to one dimension requires a corresponding reduction in attention to others. The parameter works in the same way as equivalent parameters in models of categorization: The distance between any pair of items along each psychological dimension is multiplied by the attentional weighting parameter for that dimension before the distances along the different dimensions are summed to enter into the similarity computation.

Thus, in the two-dimensional version of SIMPLE there are four free parameters: c, t, s, and w_T. When the model is applied to the Underwood (1957) data, the pattern shown in Figure 7 (dashed line, labeled *Model Two*) is obtained. Average item recall probability was used as the performance measure. Again it can be seen that performance drops off substantially as the amount of interference from previous lists increases.

In summary, SIMPLE shows PI, which increases as a function of the number of previous lists. PI is observed whether lists are modeled as single items or as lists. The PI arises from the reduced discriminability of items when they have increased numbers of nearby temporal neighbors. We have also introduced the idea of a two-dimensional space in which items may be located, to capture the idea that items' positions are represented on both a within-list nontemporal dimension and a more global temporal dimension; we return to this theme below in the context of grouping effects.

Forgetting Over Time: Peterson and Peterson (1959)

The simulations above illustrate forgetting due to PI after a retention interval that is long in relation to the time assumed to intervene between interfering and target items. Do similar principles apply in the case of short-term forgetting? Such questions are central to the scale-similar memory assumption. The rapid forgetting over time of consonant pairs (J. Brown, 1958) or trigrams (Peterson & Peterson, 1959) when rehearsal was prevented during the retention interval was interpreted at the time as evidence for trace decay models (J. Brown, 1958) and as evidence that different principles governed short- and long-term forgetting. If correct, such an interpretation would be problematic for scale-independence claims. We therefore investigated the ability of SIMPLE to account for Brown–Peterson forgetting, using the data of Peterson and Peterson (1959). (The J. Brown [1958] data are less suitable for modeling, as a smaller number of retention intervals was used.) The data (correct recalls with latencies below 2.83 s) are reproduced in Figure 9.

We reproduced the Peterson and Peterson (1959) experimental conditions as closely as possible, treating each consonant trigram as a single item. Recall of the entire trigram was assumed to be the cube of the probability of recalling a single item (i.e., independent recalls were assumed, although this is undoubtedly an oversimplification; Schweickert, Chen, & Poirier, 1999).

In modeling the Peterson and Peterson (1959) data, a gap of 15 s between trials was assumed (as in the experiment). We assumed in the simulation that 10 previous items had been presented before the critical item, to allow for PI. Retention intervals of 1 through 18 s were used, as in the experiment, and we assumed a response time (additional to the stated retention interval) of 1 s. The results are shown in Figure 9; a reasonable fit was obtained. The forgetting curve produced by the model is quite well approximated by an exponential function; we return to this issue later. As Laming (1992) noted, the possibility of covert rehearsals may lead to shorter effective retention intervals than those measured experimentally. Laming's model gives a better fit under this assumption, and similarly, a slightly better fit may be found with the present model if we were to subtract a constant from all retention intervals.

Reduced Forgetting in the Absence of PI: Keppel and Underwood (1962)

Peterson and Peterson (1959) specifically argued against interference-based explanations of their results. However, the simulation above showed that Brown–Peterson forgetting can occur as a result of interference in the SIMPLE model. This is consistent with the classic data of Keppel and Underwood (1962), who found that on the first trial of a typical short-term memory experiment, there was little or no forgetting over time. On the second and subsequent trials, in contrast, memory performance was worse at longer intervals and the rate at which it reduces was dependent on

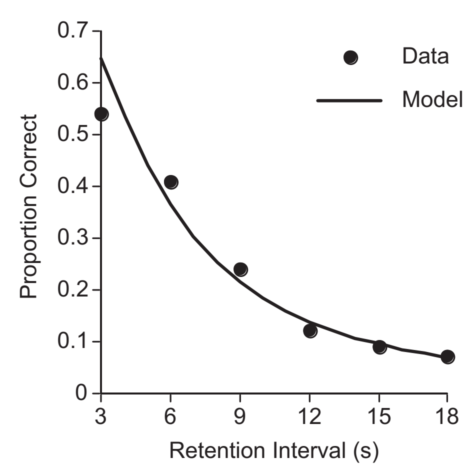

Figure 9. Form of the forgetting function between 3 and 18 s. Filled circles show data (adapted from Peterson & Peterson, 1959); solid line shows behavior of SIMPLE (scale-independent memory, perception, and learning model).

the number of previous trials (there was faster forgetting with a larger number of previous lists). The basic pattern of results found by Keppel and Underwood is reproduced in Figure 10A. (Note, however, that in some of their other studies Keppel and Underwood did find rather more forgetting on the first trial.)

This study used trigrams as stimuli; these were treated as in the previous demonstration. The behavior of a one-parameter model (not incorporating a threshold) is shown in Figure 10B. We assumed a 60-s gap between trigrams and that "immediate" recall took place after 1 s. The model reproduces the main features of the data well. The behavior of the model can be explained in terms of the same principles as before. First, it predicts no forgetting on the first trial, regardless of the retention interval, because there is no competition from previous list items. Second, the model predicts faster forgetting when there are more previous lists. This is because after a long retention interval, the effect of Weberian compression is such that the previous item becomes closer on the scale to the current item. This means that a temporal cue is more likely to lead to retrieval of the prior item than at shorter intervals. This effect of Weberian compression will be greater when there are more previous trials, and this is the reason for the greater rate of forgetting after a larger number of previous trials in the model. Forgetting is a little faster in the model than in the data. This may reflect the fact that recall of each of the three consonants making up a trigram is not completely independent, shorter actual than intended effective retention intervals due to covert rehearsal (cf. Laming, 1992), or the absence of a threshold in the one-parameter model. Inclusion of a threshold leads to a slightly better fit (Figure 10C). In summary, the behavior of the model is consistent with the

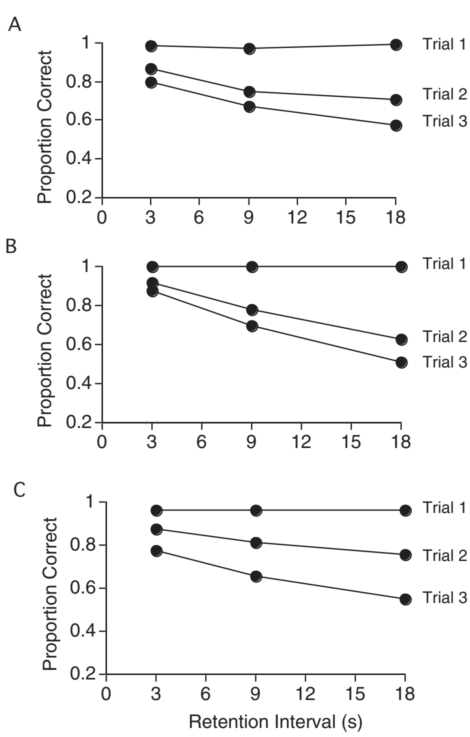

Figure 10. A: Proportion correctly recalled in a Brown–Peterson task as a function of trial number and retention interval (data adapted from Keppel & Underwood, 1962). B: Behavior of the one-parameter model. C: Behavior of the three-parameter model.

conclusions drawn by Keppel and Underwood (1962) themselves: Interference may provide an account of short-term Brown–Peterson forgetting in terms of the same interference mechanisms that may explain forgetting over much longer timescales.

Further Evidence Against Trace Decay: Turvey et al. (1970)

This explanation of PI as a fundamental source of forgetting also explains the striking data reported by Turvey et al. (1970). In a modified Brown–Peterson task, Turvey et al. varied, between participants, the duration of the filled retention intervals. For the first four trials, one group had 10 s, a second group had 15 s, and a third group had 20 s of distraction prior to recall.[11] All groups had equivalent buildup of PI by Trial 4. On the fifth trial, however, all groups counted backward for a retention interval of 15 s. Of most interest is the finding that performance for the 10-15 group became worse, that for the 15-15 group remained approximately constant, and that for the 20-15 group increased. The result is shown in Figure 11. Thus, performance is determined not by retention interval (which was the same for all three groups) but by the interaction between (a) retention interval and (b) the temporal distance between the item to be recalled and proactively interfering items. These counterintuitive results represent a clear challenge to simple trace decay accounts and have been taken as evidence for temporal trace-discriminability ratiolike accounts (e.g., Baddeley, 1976; Neath, 1998). Can SIMPLE account for them? We used the same method as used to investigate Brown–Peterson forgetting in the simulations above, although without assuming previous lists. With c as the only free parameter, the results shown in Figure 11 were obtained.

Why does SIMPLE behave in this way? Consider first the case where retention interval remains constant within a condition, as in the first four trials of the experiment. With a short retention interval, only a small amount of time elapses between presentation and recall, and so there is relatively little Weberian compression of the temporal location code for the to-be-remembered item. This should lead to good performance. On the other hand, the immediately prior item was learned only a short amount of time previously, and so will cause interference at retrieval. When the retention interval is longer, there will be a greater amount of elapsed time and more Weberian compression of the to-be-remembered item. This should lead to worse recall for the item as compared with the shorter retention interval. Set against this, the longer retention interval also applied to the previous trial, and so there is less interference (according to the local distinctiveness principle) from previous trials than in the short retention interval comparison. These factors act to oppose each other, with the net result that performance on the fourth trial of the experimental series is similar for all retention intervals. (This account differs from that of other temporal-discriminability memory models in that we assume here that all previous items enter into the calculation of memorial discriminability.)

[11] Two other groups were included: a 5-15 group (where the decrease in memory performance on the final trial was similar to that seen in the 10-15 group) and a 25-15 group (where the increase in memory performance on the final trial was actually somewhat larger than that seen in the 20-15 group shown).

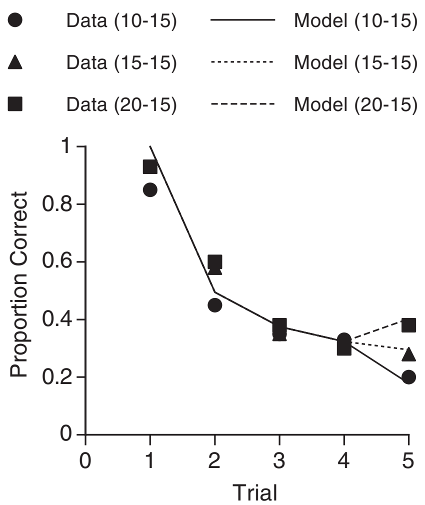

Figure 11. Proportion correctly recalled in a Brown–Peterson task as a function of intertrial spacing and retention interval (data adapted from Turvey et al., 1970). Lines show the performance of SIMPLE (scale-independent memory, perception, and learning model).

On the fifth trial, when the retention interval either increases (the 10-15 condition) or decreases (the 20-15 condition), the conditions separate. In the 10-15 condition, the longer retention interval worsens performance relative to the previous (fourth) trial, and the relative closeness of the interfering trials (approximately 30, 20, and 10 s prior to the target item) also depresses performance. In the 20-15 condition, in contrast, less Weberian compression occurs in the 15-s retention interval than occurred over the 20-s retention interval of the previous trial, but the interfering trials are still temporally distant (60, 40, and 20 s away). Performance therefore improves on the final trial.

Trace decay models of short-term memory have continued to receive attention despite the findings of Keppel and Underwood, Turvey et al., and others. This continued attention arises partly because of the success of trace-decay/rehearsal models in accounting for item duration effects (e.g., Baddeley, Thomson, & Buchanan, 1975; but see, e.g., G. D. A. Brown & Hulme, 1995; Neath, Bireta, & Surprenant, 2003) and partly because of the lack of an adequate interference-based account of the relevant short-term memory phenomena (i.e., an account of the type we are attempting to develop here). However, an additional reason relates to the findings of Baddeley and Scott (1971; see also Marcer, 1972) that forgetting can occur following the first presentation of a list, and these findings continue to be interpreted as evidence for trace decay.

Within a model like SIMPLE, forgetting of items within a single sequence can occur owing to the difficulty of temporal discrimination of the separate items *within* a sequence. Baddeley and Scott (1971) excluded this kind of intrasequence explanation of their observed forgetting (Melton, 1963) on the grounds that such an account would incorrectly predict faster forgetting for longer sequences. However, a model like SIMPLE need not necessarily produce faster forgetting for longer lists, because the amount of interference for a given item, being determined mainly by near neighbors, is not necessarily much influenced by additional items in a longer list (in practice the behavior tends to be parameter and paradigm dependent). Thus, forgetting of a single list of items in the absence of proactively interfering material need not be interpreted as evidence for trace decay. Intrasequence interference can potentially explain single-trial forgetting, without making incorrect predictions about the rate of forgetting for longer lists.

Release From PI: Loess and Waugh (1967) and D. D. Wickens, Born, and Allen (1963)

Key evidence in favor of PI explanations of short-term forgetting is the *release-from-PI* phenomenon. When several trials of a given type (e.g., letter trigrams) are used in a Brown–Peterson paradigm, PI builds up and recall reduces across trials. Performance on a subsequent trial may improve dramatically when the nature of the material is changed (e.g., from letters to digits), and this is assumed to be due to the reduction in interference from previous trials as they are no longer similar to the target item (e.g., D. D. Wickens, Born, & Allen, 1963). The greater the change in the nature of the material is, the greater is the release from PI. The effect appears to be due to retrieval rather than encoding processes (Gardiner, Craik, & Birtwistle, 1972). Release from PI also occurs when the time interval separating successive to-be-recalled items is increased sufficiently (Loess & Waugh, 1967; Peterson & Gentile, 1965), consistent with the idea that the temporal-discrimination problem becomes easier under such circumstances (Baddeley, 1976; Baddeley & Scott, 1971).

Because there is a large literature on release from PI, we focused on SIMPLE's ability to explain just three basic phenomena. These are (a) the release from PI after a shift in the nature of the material, (b) the dependence of the size of the release on the amount of change in the to-be-remembered material, and (c) the release from PI due to the passage of time alone. In all cases it is possible to explain the phenomena using the same mechanisms of discrimination based on distinctiveness within a local psychological neighborhood. Items will be easy to discriminate just to the extent that they are isolated from their near neighbors in psychological space. When both temporal neighborhood and semantic neighborhood are relevant, as in the classic release-from-PI paradigm, the discriminability of an item in memory will depend both on its near neighbors in semantic space and on its near neighbors in temporal space, and so the model must be extended to incorporate semantic as well as temporal neighborhoods.

In some simulations above we extended the relevant psychological space from a unidimensional time line to a two-dimensional space, where the two dimensions represented list within trial and item within list. To model release from PI, we again need to assume a two-dimensional space within which local distinctiveness is calculated, but here one dimension will represent temporal distance whereas the other will represent item similarity. A simple space is illustrated in Figure 12, for the case where three consonant trigrams are presented followed by one three-digit item.

The horizontal axis dimension represents temporal position; items are assumed to be presented at regular temporal intervals, and so in psychological space they are logarithmically spaced along this dimension as in previous simulations. The vertical axis represents item similarity. The consonant trigrams are assumed to be similar along this dimension and are assigned arbitrary but equal values (all items in Figure 12A; first three items in Figure 12B). The group of digits are, in contrast, assumed to have a distinct value along this dimension (see fourth item in Figure 12B)

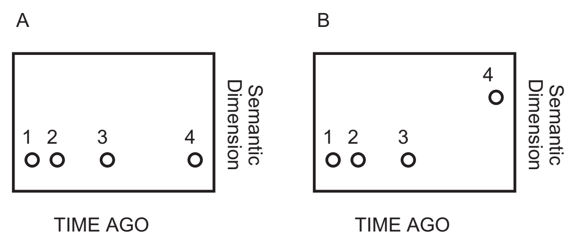

Figure 12. Illustration of how release from proactive interference may result from local distinctiveness principles. A: All items from same semantic category. B: Items 1–3 from same semantic category; Item 4 from different semantic category.

and are assigned a different value along the "semantic" dimension. In the two-dimensional space, they are therefore in much less dense local neighborhoods than the third consonant trigram, and this would be expected to lead to superior memory performance. The use of a single dimension to represent item similarity is, of course, a simplification. However, it suffices for illustration, for it can be thought of as representing a conceptual quality such as "digitlikeness." Whereas it would in principle be possible to construct a much richer multidimensional semantic space, as is typically done in exemplar models of categorization, this is not necessary for present purposes. Our point is simply that a change in item type can result in enhanced local distinctiveness, and hence memorability, for that item even when it is part of a temporally regular sequence.

We examined the behavior of SIMPLE under the conditions described. Four items were presented to the model, separated by 10 s. The first three items were given values of 1 on the "conceptual" dimension, whereas the fourth item was given a value of either 1, 1.5, or 2.5 (to represent no category shift or a small or large category shift). An effective retention interval of 2 s was assumed, and as before, recall probability was calculated as the cube of the probability of recalling a single item. For simplicity, the no-threshold version of the model was used. The results are shown in Figure 13A, where the classic data pattern is produced: Performance reduces across the first three items, due to a buildup of PI, and then increases dramatically on the fourth item, illustrating release from PI. Furthermore, the extent of the release from PI is dependent on the extent of the category shift, as in the data (Gardiner et al., 1972; D. D. Wickens et al., 1963).

We also examined release from PI as a function of the passage of time alone (Loess & Waugh, 1967; Peterson & Gentile, 1965). As a number of authors have pointed out, release from PI after a temporal gap is difficult for many forms of classic interference theory to predict, because such accounts would expect spontaneous recovery of interfering items as a function of the passage of time, and so there should be no reason to expect any time-based release from PI. We examined memory for 12 trigrams, arranged into three blocks of four items. Trials within blocks were separated by 1 s, and blocks were separated by 120 s. With other assumptions as in the previous simulation, the results shown in Figure 13B were obtained. Again, the right general pattern was obtained, with a gradual buildup of PI when items follow in quick succession and release from PI after a longer interval (Loess & Waugh, 1967). The explanation of the model's behavior is similar to those that have been given before: Given a temporal retrieval cue for a given item, that item will be more retrievable to the extent that it is locally distinctive in its temporal neighborhood. Items that have been closely preceded by two other items will be much less temporally distinctive than will items preceded by a large temporal gap, and so memory is better for these latter items. This general behavior of the model—time-based release from PI—is consistent with a considerable amount of evidence that PI is reduced by temporal separation both in AB-AD paradigms (Keppel, 1964; Underwood & Ekstrand, 1967; Underwood & Freund, 1968) and over shorter time periods (Alin, 1968; Kincaid & Wickens, 1970; Peterson & Gentile, 1965). Note that SIMPLE offers essentially the same explanation of time-based release from PI as was given for the ratio-rule-like phenomena and for the Turvey et al. (1970) data. The account of spontaneous recovery offered by Estes (1955) resonates well with the account here (see also Mensink & Raaijmakers, 1988).

The Time Course of Forgetting

We are now in a position to examine the time course of forgetting in the model. This is a question of central theoretical importance; any plausible model of memory must surely have something to say on the form of, as well as the fact of, forgetting. As noted in the introduction, the possibility that the time course of forgetting may approximately follow a power law is of particular interest in view of the scale independence of memory retrieval principles and

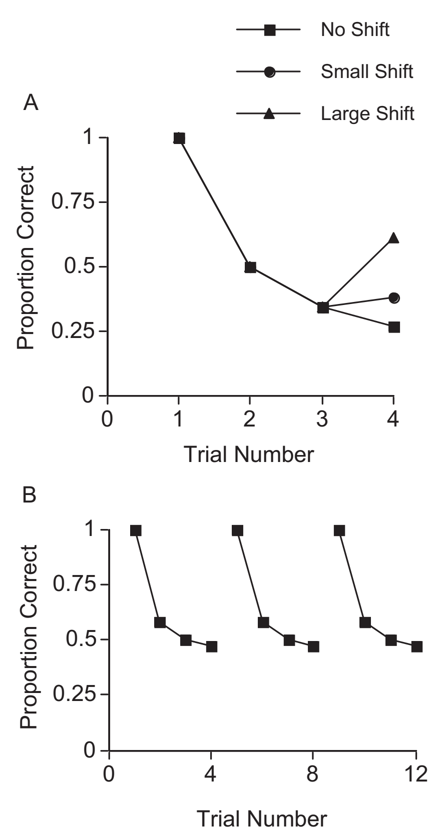

Figure 13. A: Release from proactive interference in SIMPLE (scale-independent memory, perception, and learning model) as a function of the size of the category shift. B: Release from proactive interference in SIMPLE as a function of the passage of time.

adaptive considerations (Anderson & Milson, 1989; Anderson & Schooler, 1991). A power function has the form $P = aT^{-b}$, where P is the measure of memory performance, T is time elapsed, and a and b are constants.

We address three related questions: (a) Does the SIMPLE model necessarily exhibit precise power-law forgetting, (b) is forgetting in the model generally well characterized by a power-law or some other function, and (c) how variable is the form of the forgetting curve produced by the model? Of particular interest is the possibility that small and theoretically insignificant parametric alterations might change the form of the forgetting curve. If so, the form of the forgetting curve may not provide a naturally invariant characteristic of memory. To anticipate our conclusion: SIMPLE does not predict any simple form of forgetting curve, because the form of the forgetting curve will depend on methodological details of a given procedure and, in particular, on the extent of PI from other items. The results of exploration with SIMPLE therefore raise the possibility that the difficulty in obtaining simple universal forgetting functions, despite a century of effort (Rubin & Wenzel, 1996), may be a natural consequence of any model likely to provide an account of PI effects, and that the form of the forgetting curve does not necessarily provide a psychologically useful level of description.

First we note that despite the ratiolike properties of SIMPLE, power-law forgetting curves need not automatically follow (this can be seen by considering retrieval probabilities for a simplified case where there are only two items in memory). It is important to note, however, that such a result does not exclude the possibility that when a more realistic experimental situation is modeled, a good approximation to power-law forgetting will be obtained. The form of the forgetting function seems likely to depend on the spacing of other competing items in any model that accounts for effects of the type we have modeled above.

We therefore turned to our second question and examined the time course of forgetting in the model using the methodology described in previous demonstrations above for two-dimensional memory representations (the two dimensions being temporal distance and within-list position). First, we examined recall of a five-item list, preceded by four previous lists (to allow for a realistic amount of PI). Two seconds separated each item within a list, and 12 s separated each list. We examined forgetting for retention intervals of between 2 and 100 s (the shortest retention intervals were not used, to avoid possible artifacts due to ceiling effects and the use of a retention measure with a maximum value of 1), for three values of c. The resulting forgetting curves are shown in Figure 14, together with two-parameter power-law fits. Good fits were obtained for all parameter values. Lower R^2 values were obtained for simple exponential and logarithmic functions (mean R^2 values of .83 and .97, respectively). However, we did not engage in extensive curve fitting or model comparison (cf. Myung & Pitt, 1997) as our aim was merely to examine whether SIMPLE would exhibit a reasonable approximation to power-law forgetting. In a series of unreported simulations we observed that a good power-law fit was also obtained for different values of w_T, although logarithmic curves sometimes fitted performance as well as did power-law curves.

The choice of performance measure is important in curve-fitting exercises of this type (see Rubin et al., 1999; Rubin & Wenzel, 1996; T. D. Wickens, 1999, for recent discussions). Power-law curves must predict infinite performance at zero retention intervals, but performance below 100% cannot be observed on a "proportion correct" performance measure. Some authors therefore use "recall odds" ($p/[1 - p]$, where p is the probability of recall) as the performance measure (see, e.g., Anderson & Schooler, 1991). When recall odds are used as the measure of performance in SIMPLE, using the same simulated experimental conditions as described above, power-law curves fit much better than exponential or logarithmic curves, even when all retention intervals from 1 s up to 250 s are used. (For example, for the data in Figure 14, the mean R^2 was .99, .86, and .73 for power, logarithmic, and exponential functions, respectively.) The success of power-law fits in this case is attributable partly to the power law's particular suitability for capturing the very rapid rise in recall odds as the performance measure approaches infinity as the retention interval tends toward zero. However, odds measures can be unstable at low retention intervals, and reliable empirical data are difficult to obtain.

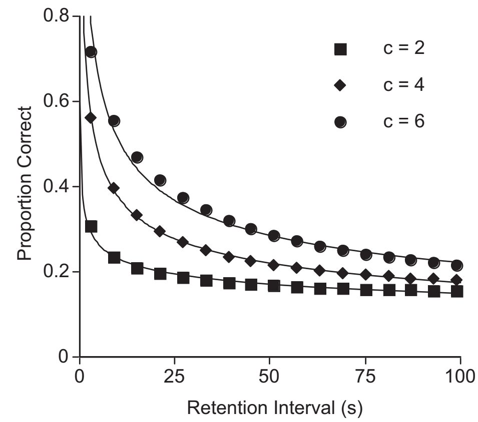

Figure 14. Form of the forgetting function produced by SIMPLE (scale-independent memory, perception, and learning model) as a function of the c parameter. Solid lines show best fitting power-law curve.

We therefore return to the use of proportion correct as the performance measure, as this has been most widely adopted in the empirical literature, and examine the possibility that different curves might fit SIMPLE's forgetting performance over different timescales. Such a finding would be of some theoretical interest, for different forms of forgetting over different timescales have sometimes been taken as evidence for the operation of different memory stores. The fit of the SIMPLE model to the data obtained by Peterson and Peterson (1959; Figure 9 above), which pertain to forgetting over 18 s, is replotted in Figure 15. Panel A plots the output of the model on logarithmic axes; this axis transformation will lead to a power-law curve appearing as a straight line. It can be seen that the behavior of the model systematically deviates from the best fitting power-law curve. Panel B plots the same data with logarithmic transformation of the y-axis only. This axis transformation will lead to an exponential curve appearing as a straight line. It is evident that an equally good fit to the model's forgetting in Peterson and Peterson's (1959) paradigm is given by an exponential curve ($R^2 = .99$); better fits can be obtained if the simulation is run specifically to obtain them.

Thus, taking all of the results together and considering just logarithmic, exponential, and power functions, forgetting in the SIMPLE model closely follows a power law when range artifacts

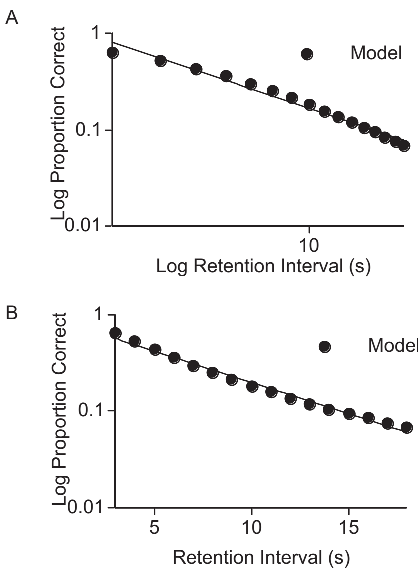

Figure 15. Forgetting curve of SIMPLE (scale-independent memory, perception, and learning model) over short time interval fit by power law (A) or exponential (B) curve.

are avoided. However, under particular circumstances forgetting in the model may be best described in terms of a power law (longer timescales; use of recall odds as performance measure), a logarithmic function (when no attention is paid to within-list distinctiveness), or an exponential function (short timescales when proportion correct is the performance measure). Further findings could be reported if we were to consider the 105 functions considered by Rubin and Wenzel (1996), all of which were rejected on the basis of new data by Rubin et al. (1999; see also Chechile, 2006). Yet exactly the same memory model is being used in all cases. We therefore endorse the conclusion of T. D. Wickens (1999): It seems unlikely that a single form of forgetting curve will apply across different methodological circumstances. We can add a demonstration that a single relatively simple architecture can, under different circumstances, instantiate a variety of forgetting functions. An example is provided in Figure 16, which shows that forgetting (five-item list; case described above with $c = 6$) is very well characterized by an exponential curve for the first 15 s of retention ($R^2 = .99$) and a power law thereafter ($R^2 = .999$). Clearly it would be wrong to conclude from this that there are two separate stores with different characteristics that operate over different timescales.

Summary of Section Results

What unites the simulations we have reported in this section is the emphasis on the interactive roles of PI and the passage of time. One key conclusion is that the appearance of forgetting due to the passage of time alone can result from such a model as SIMPLE because the relative importance of even a single proactively interfering item will become progressively greater as time passes, due to Weberian compression of the scale on which both target and interfering items are represented. Thus, the model produces some of the results that have traditionally been used to support time-based decay even though there is no decay—time based or otherwise—anywhere in the model. A second key conclusion is that different functional forms of forgetting curves over different timescales may be found despite the absence of a separate short-term memory system.

Series 3: Serial Recall

Whereas the previous simulations focus on time and PI, they have had little to say about serial recall. Can the same framework be used to examine both serial and free recall? In this final set of simulations we focus on serial recall, with a specific focus on (a) serial position curves and error movement gradients in serial recall and (b) effects of phonemic confusability and their interaction with retention interval. The key theoretical focus remains the same: to explore the possibility that the same principles govern retrieval over many different timescales and the possibility that several short-term memory data can be explained without an assumption of trace decay. An additional concern is whether the differences between serial and free recall can be understood within a unitary framework in terms of differing task requirements.

Error Movement Gradients I: Nairne (1992)

A central characteristic of SIMPLE is that items near to one another in psychological space will be confusable. The tendency for systematic order errors to occur has been well documented over the past quarter century. In serial recall tasks and order reconstruction tasks, the same basic effect is consistently found: Items that are not recalled in their correct serial position are most likely to be recalled in a serial position adjacent to the correct one, and are progressively less likely to be recalled in a position away from the correct one as the distance between target (correct) position and recalled position increases (see, e.g., Estes, 1972; Healy, 1974; Henson, Norris, Page, & Baddeley, 1996; Nairne, 1991, 1992).

We first consider the simple case where participants are presented with a list of items in serial order and then at recall are given the items and asked to arrange them in the order in which they were presented. In the first demonstration we ignore effects of prior lists and the time-course of recall. Figure 17 shows the results of an experiment of this type conducted by Nairne (1992). Participants were presented with five lists of five items and required to rate them for pleasantness; at test, participants were provided with

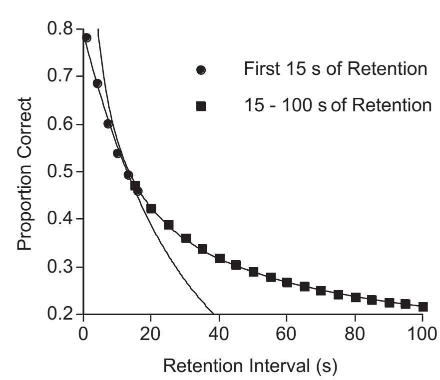

Figure 16. The different form of forgetting function shown by SIMPLE (scale-independent memory, perception, and learning model) over short and long retention intervals.

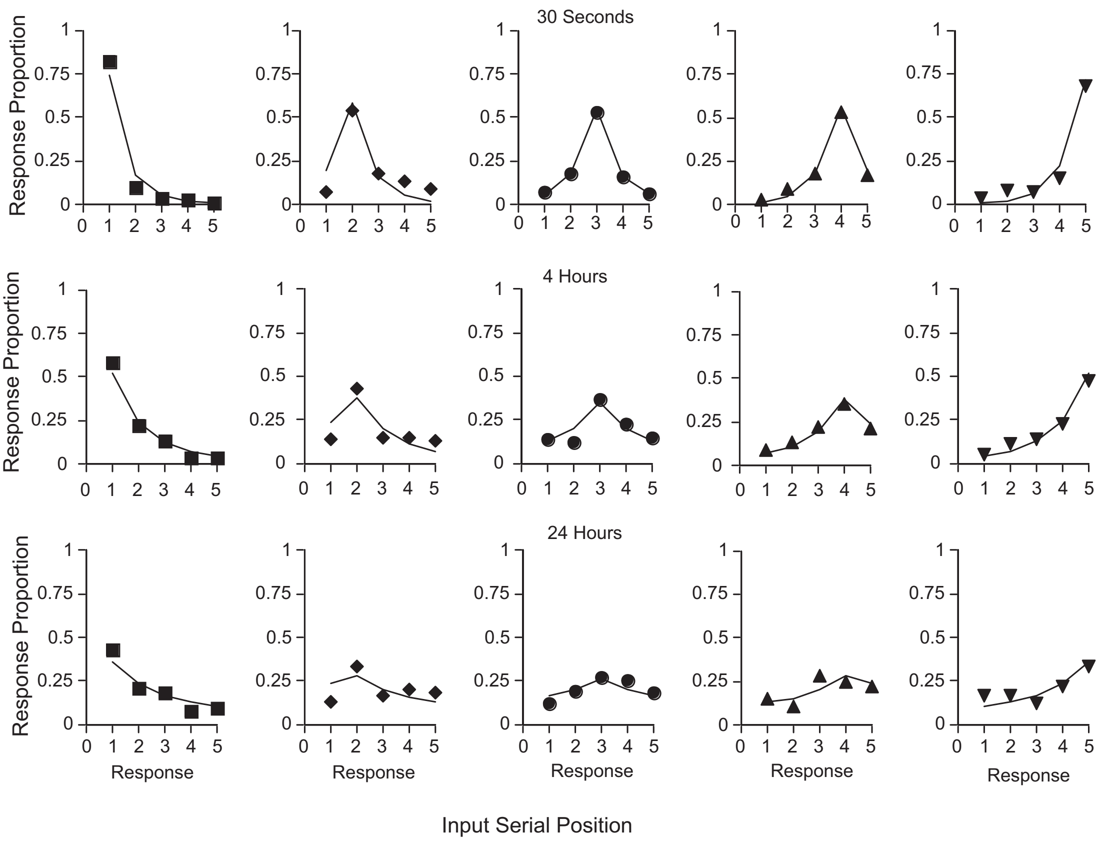

Figure 17. Positional uncertainty gradients after retention intervals of 30 s, 4 hr, or 24 hr (data adapted from Nairne, 1992). Lines show the performance of SIMPLE (scale-independent memory, perception, and learning model).

the items from each list and had to place them in the order of presentation. The positional uncertainty gradients represent participants' responses after 30 s (Panel A), 4 hr (Panel B), and 24 hr (Panel C); each panel shows the different output positions into which an item was placed.

In addressing these data, we need add no new assumptions to the model. We simply calculate the probabilities of recall for every item, including the correct item. This is the same method used in all previous simulations. The new aspect is the additional calculation of the probability of (incorrectly) recalling each item in a nontarget position. Figure 17 shows the positional uncertainty gradients produced by the model for the three levels of retention interval. Following Nairne (1992), we assumed 2.5 s between items and retention intervals of 30 s, 4 hr, or 24 hr. However, the input to the model was simpler than that used in the experiment, for we assumed only a single five-item list (cf. five five-item lists in the experiment).

It can be seen that the model produces the characteristic positional uncertainty gradients as shown in Figure 17. The values of c were allowed to vary with retention interval; there were thus just three values of one parameter for the 75 data points. The parameter values illustrate the strong tendency for longer retention intervals to be associated with higher c values; we return to the issue of cross-simulation parameter values in the General Discussion. As omissions are not possible in order reconstruction tasks, no threshold mechanism is needed, but any response bias was removed from the model's output using iterative normalization (because the experimental procedure forces each response to be produced equally often). The model gives a reasonable account of error movement gradients in an order reconstruction task and the qualitative similarity of such gradients over a wide range of retention intervals. Although the R^2 value is low for the longest retention interval, the qualitative effects are captured. The basic effect—that the probability of placing an item close to its correct position is greater than the probability of placing an item far from its correct position—follows straightforwardly from the architecture of the model. The similarity of the memory codes for any two temporal positions falls off as a negative exponential function of the temporal distance between them, and this similarity is reflected in the movement gradients.

One key piece of evidence for hierarchical models of memory for serial order (e.g., G. D. A. Brown et al., 2000; Estes, 1972; Henson, 1998b) is the observation that similar error movement gradients can be seen at the level of lists within a trial and at the level of items within a list (Underwood, 1977). For example, Nairne (1991) presented five lists of five items and asked participants to rate items for pleasantness. Two minutes later, he presented five lists of five blanks each and also the list of 25 words; participants were asked to place the words in their original list and within-list position. The characteristic uncertainty gradients were seen for both the list and the within-list dimensions. These data are shown in Figure 18A and 18C.

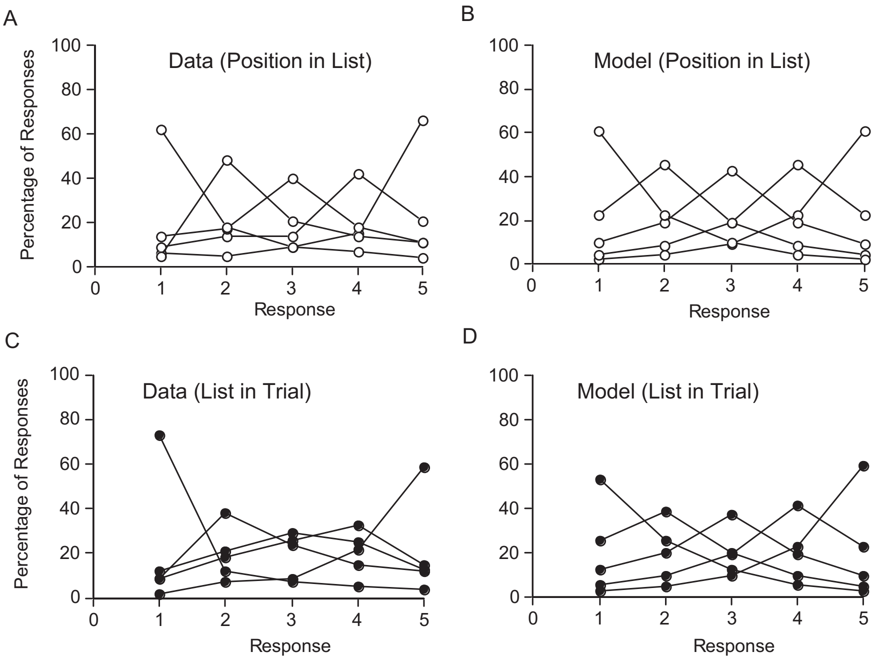

Figure 18. Positional uncertainty gradients for the dimensions of within-list position (A) and list-within trial (C); data calculated from Nairne (1991). B and D: Performance of SIMPLE (scale-independent memory, perception, and learning model).

SIMPLE can be extended to represent the position of an item on two (or more) hierarchically arranged dimensions simultaneously. This was done for the two-dimensional model introduced in the discussion of the Underwood (1957) findings above and the model of the release from PI. There, we assumed that items could be viewed as located in a two-dimensional space, where one dimension represented the position of an item within the trial as a whole (i.e., the set of 25 words in the Nairne [1991] experiment) whereas the other dimension represented the position of an item within each five-item list. This scheme was illustrated in Figure 8 for shorter and fewer lists, and we applied it here to the Nairne (1991) results.

Following the Nairne (1991) methodology, in the model it was assumed that items were separated from one another by 2.5 s and that lists were separated from one another by 5 s. A retention interval of 120 s was added in both cases. The items were represented in a two-dimensional space, with one dimension corresponding to within-list position and one dimension corresponding to within-trial position. There were just two free parameters to predict the 625 possible data points (the probability of recall in each of the possible 25 output positions for each of the 25 items): c and w_T, the attentional weight given to the within-trial temporal distance dimension. Model response bias was removed as before. The resulting positional uncertainty gradients are shown in Figure 18B and 18D, where it is evident that the main features of the data are captured by the simple two-parameter model. Nairne (1991) found a small degree of nonindependence in his data (i.e., the probability that an item would be placed in the correct position on one dimension was not independent of the probability of correct positioning on the other dimension). Additional assumptions would be needed to capture this nonindependence in SIMPLE. For example, if some items are not properly registered at encoding, such items are likely to be placed in incorrect positions on both dimensions at recall, leading to nonindependence.

This simulation has extended the scope of the model to account for the positional uncertainty gradients that are observed over two different timescales simultaneously; the account is achieved by assuming that items are located, and retrieved, on the basis of their positions in a two-dimensional neighborhood. Application of the local neighborhood rule, in exactly the same fashion as has been done throughout, results in advantages for items located in more sparsely populated regions of psychological space. However, items that are located close to one another on the within-list position dimension may have a high probability of exchange even if they are widely separated in time, because of the two-dimensional representation within which they are located.

Effects of Acoustic Confusability in Serial Recall: Henson et al. (1996)

The effects of phonological confusability on the short-term serial recall of verbal material have been central to arguments for a separate short-term memory system. Can SIMPLE shed light on such data? We focus on accounting for two basic findings related to the effects of phonemic confusability. First, and most basic, is the finding that short-term memory for the serial order of verbal items is reduced when the items are phonologically confusable, with the additional errors being predominantly movement errors (transpositions; see Baddeley & Ecob, 1970; Conrad, 1964, 1967; Estes, 1973; Healy, 1975). Second, we examine the tendency for items recalled in the wrong serial position nevertheless to be recalled in positions close to the correct one. An additional question is whether extended primacy and minimal recency, which are characteristic of forward serial recall, will emerge in the model as a consequence of forgetting over time during the recall process.

In all cases, we assume that confusability effects can be accommodated in terms of an extra dimension in the psychological space within which items must be discriminated in a memory task. This is illustrated, for the case of a single list of six items, in Figure 19. The horizontal dimension represents the temporal distance of items at the time list presentation is complete, as normal. The vertical dimension represents a "confusability" dimension, such that items with similar values on this dimension have similar phonological representations, whereas dissimilar items have more widely spaced values on this dimension.

Figure 19A depicts the case where no items are very confusable—the items are evenly spaced during presentation and are phonologically distant from one another. Figure 19B shows the case where the items are more phonologically similar—they are therefore distinguishable from one another primarily in terms of their position along the temporal dimension but have similar values on the second, phonological dimension. Simply by examining the items' local neighborhoods in this two-dimensional space, we can see that dissimilar items will be better remembered than similar items because the dissimilar items have fewer near neighbors.

To illustrate, we examined SIMPLE's memory for a single list of six items, separated by 0.4 s, at immediate recall (implemented as a delay of 0.5 s before recall of the first item). To allow for the possibility of a realistic amount of PI we assumed the presence of four previous lists, with each list separated by 30 s. Typical results from an experiment of this type (Henson et al., 1996) are shown in Figure 20 in Panels A and C, which illustrate the characteristic pattern of better overall performance for nonconfusable than for confusable lists, with the additional movement errors for the confusable items. The simulated experimental conditions simulated correspond closely to those adopted by Henson et al. (1996).

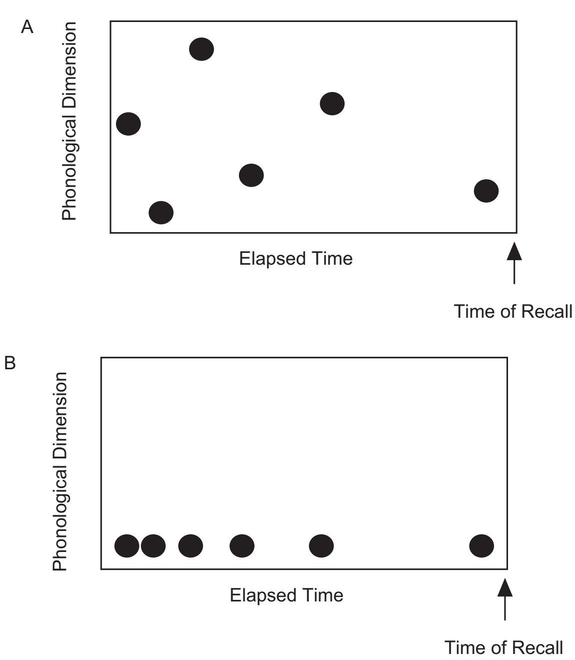

Figure 19. The representation of a six-item list along both temporal distance and phonological confusability dimensions. A: Dissimilar items. B: Similar items.

In the model, the position of items was represented in the normal way—that is, in terms of their temporal distance from the point of recall. Because precise output timings are not known, we assumed a linear rate of output, with each additional item taking 1 s to recall. These figures can in principle be set independently by controlling or measuring output recall times (Lewandowsky, Brown, Wright, & Nimmo, 2006; Lewandowsky, Duncan, & Brown, 2004; Surprenant, Neath, & Brown, 2006); in practice, good model fits were achievable under a range of assumptions about the time course of output (although see below). In addition to their values on the temporal dimension, items were assigned identical or different values along a phonological dimension according to whether lists were composed of similar or dissimilar items, as illustrated in Figure 19. Items in a "similar" list were assigned identical numbers on this dimension (one of the values 1 though 5); each item in a "dissimilar" list was randomly assigned one of the values 1 though 6 without replacement (50 different random assignments were used). As in previous demonstrations an attentional weight parameter, w_T, was used to specify the relative weight given to the temporal dimension over the phonological dimension.

The results are shown in Figure 20B and 20D. Omissions are possible and observed in serial recall, and there are therefore four parameters, c, s, t, and w_T, that could not in principle be set from knowledge of the experimental conditions. A good fit to the 72 data points was obtained, with all parameters except for t being constant across conditions.

In summary, the SIMPLE model can be extended to account for confusability effects in memory. This is done by extending the dimensionality of the space in which memory items are stored, so that one dimension represents temporal position and the other represents degree of confusability. In principle, it would be possible to use independently derived metrics of acoustic confusability (e.g., Miller & Nicely, 1955) to determine the positions of items in a multidimensional phonological space. Surprenant et al. (2006) adopt a related approach to account for age differences in serial recall within SIMPLE.

We have focused on the simplest possible explanation of phonological confusability effects in serial recall in terms of the proximity of items' episodic traces in a two-dimensional space. Additional parameters, such as noisy output thresholds, can be incorporated to account for additional factors, such as the small number of observed omission errors and repetition omissions; the model behaves in plausible and predictable ways when such refinements are incorporated, exhibiting, for example, a tendency for more omission errors to occur toward the end of the list. However, the relevant data and causal mechanisms are now quite well understood in the context of previous models, and so we do not repeat previous theoretical work here. The approach can also offer a perspective on alternating list effects, whereby dissimilar items suffer little or not at all by being sandwiched between confusable items (e.g., Baddeley, 1968; Henson et al., 1996; but see Farrell & Lewandowsky, 2003). According to SIMPLE, the extent to which items from one class will benefit in serial recall from being alternated with items from another class will depend on within-class and between-class proximities in the relevant memory space; to the extent that within-class similarity is high and between-class

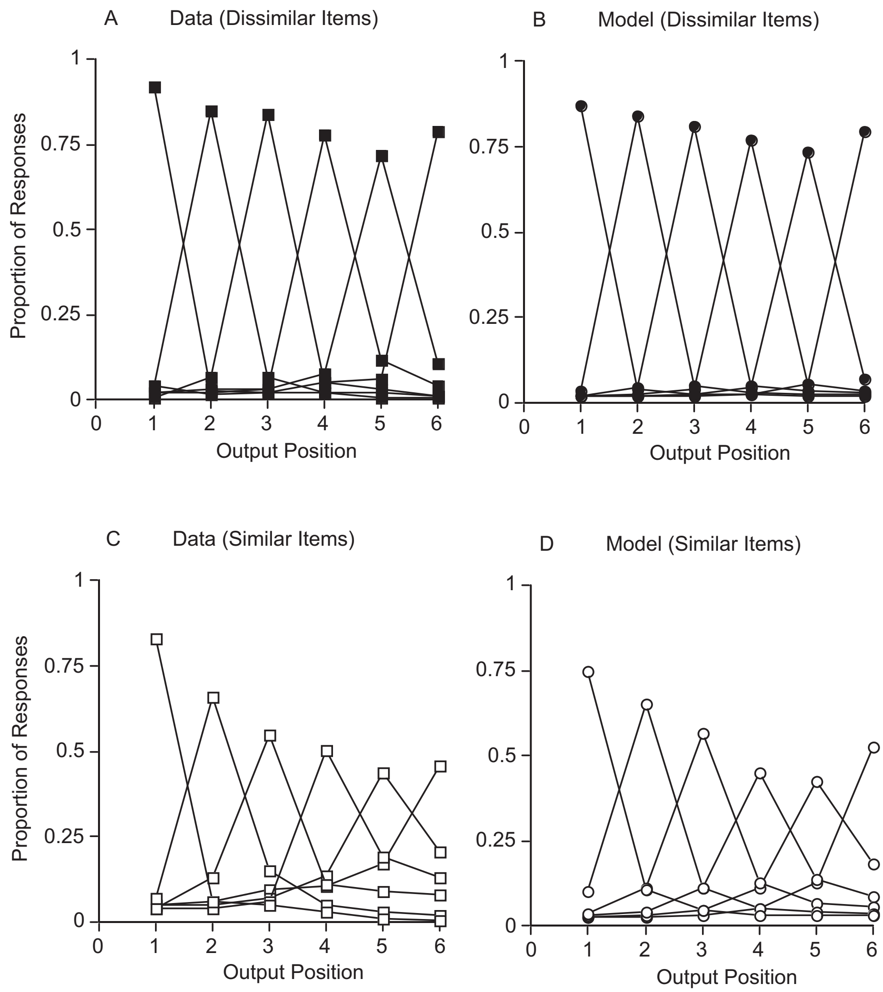

Figure 20. Positional recall functions for serial recall of dissimilar (A) or similar (C) items (data adapted from Henson et al., 1996). B and D: Performance of SIMPLE (scale-independent memory, perception, and learning model).

similarity is low, separation of items along a temporal or positional dimension will be beneficial (see also Farrell, 2006).

An important issue concerns the nature of the dimension used to represent within-list position. Because SIMPLE, like other exemplar models, assumes multidimensional memory representations, it can allow for serial order information to be represented by location on positional, temporal, or perhaps other dimensions. However, for simplicity we have assumed above that for serial recall, as with free recall, the relevant dimension is solely temporal. In the case of free recall, the assumption is consistent with the observation of recency effects and their interaction with retention interval and also fits well with the observation of temporal isolation effects (G. D. A. Brown et al., 2006). In the case of serial recall, however, extended primacy is seen and substantial recency is not typically observed when presentation is visual. The model as described above reproduces this behavior (Figure 20) only because in the simulation recall proceeds at a slower rate than list presentation. Thus, the effective retention interval (time between item presentation and item recall) is greater for late-presented (and hence late-recalled) items than for early-presented (and hence early-recalled) items. Items therefore become progressively less distinctive as recall unfolds over time, and this effect produces extended primacy. The contrast between the extended primacy seen in forward serial recall (as in the present simulation) and the extended recency seen in free recall and probed serial recall is therefore assumed to reflect the differing temporal demands of the tasks, rather than the operation of different memory retrieval principles in the different paradigms. Recency items will be superior to primacy items only if they can be recalled early, thus benefiting from their still uncompressed locations along the temporal dimension.

The small (typically 5%–10%) recency that is typically superimposed on this extended primacy in Figure 20 reflects an edge effect. However, extended primacy may also be seen in serial recall when recall is not slower than presentation. Thus, an alternative possibility is that items are represented in terms of their location along a positional (rather than or as well as a temporal) dimension and that the extended primacy partly or predominantly reflects output interference. Specifically, Lewandowsky et al. (2004) argued that when a positional (rather than temporal) dimension is used in SIMPLE to represent memory for serial order, and output interference is added into the model (producing extended primacy), a better account can be given of the effects of manipulating output time. Indeed, evidence from cross-list intrusions at the level of lists (Henson, 1999) or within-list groups (Ng & Maybery, 2002) is consistent with the suggestion that a positional dimension is used instead of (or in addition to) a temporal dimension to underpin memory serial recall (see Neath & Brown, 2007, for discussion). Recent evidence points to the near absence of temporal isolation effects in forward serial recall tasks (Lewandowsky & Brown, 2005; Lewandowsky et al., 2006), even when presentation is auditory (Nimmo & Lewandowsky, 2006) or the temporal gaps separating items extend to seconds (Nimmo & Lewandowsky, 2005). However, isolation effects, indicative of temporal encoding, may reappear even in a serial order memory task when recall is not in forward order (Lewandowsky, Nimmo, & Brown, in press). It seems likely, therefore, that attentional weight is given to a positional dimension, in addition to (or instead of) a temporal distance dimension when the task is serial recall. Although here we preserve the focus on the temporal dimension, there is no difficulty in including a positional dimension into SIMPLE's multidimensional space, as shown by Lewandowsky et al. (2004) and Lewandowsky et al. (2006). Although the addition of a positional dimension into the multidimensional space assumed by SIMPLE to underpin memory performance increases model complexity, the assumption is arguably more parsimonious than the alternative approach of assuming a completely different memory system for short-term serial recall.

Intrusions From Previous Lists

In both serial and free recall, when different items must be recalled in each list of a series of lists, many errors are intrusions from previous lists. Murdock (1974) noted that intrusions from previous lists are more likely to come from recent rather than from more distant lists—for example, an item is four times as likely to be intruded from the previous list as from the one before that. (Murdock cited this as evidence for the importance of temporal factors in recall, a conclusion that is, like much of Murdock's [1974] discussion, highly consistent with the model proposed here.) In serial recall, intrusions often occur from the recall protocol of the previous trials (see Conrad, 1960; Estes, 1991; Henson, 1996), serial position is most often preserved (see, e.g., Henson, 1998b), and fewer such intrusions occur when the temporal spacing between lists is large (Henson, 1996). Estes (1991) noted that in recall of trial n, there are rather few intrusions of items presented on trial $n - 1$ but not recalled in trial $n - 1$ (in contrast to the many intrusions on trial n that come from recall on trial $n - 1$), consistent with the suggestion that each recall is an additional learning episode capable of producing PI (Henson, 1998b).

We examined such intrusions in SIMPLE, using the account of single-list serial recall described in the previous simulation. To obtain sufficient numbers of intrusion errors from previous lists to analyze, we changed the parameter values: c was set to 2 and w_T was set to .5. (Reducing the attentional weight on the purely temporal dimension is equivalent to increasing the attentional weight on the position-within-list dimension, as the weights must sum to 1.) With these parameter settings, just over 10% of errors were intrusions from previous lists. Of these, 55% intruded into the same position as they had occupied on the previous list in which they occurred. Most intrusions came from the immediately preceding list: Of intrusion errors on trial n, 49% were items from trial $n - 1$, 23% from trial $n - 2$, and the remainder from trial $n - 3$ or earlier.

Thus, the main qualitative features observed in the data concerning intrusions from previous lists are captured, although even small changes in parameter values give rise to different quantitative data, and we did not attempt detailed quantitative fitting as suitable data sets are not available. We note that from the perspective of SIMPLE, "extraexperimental intrusions" occur via the same mechanism as intrusions from previous lists; they simply reflect PI from items beyond the temporal window of the experimental environment. From this perspective, "item errors" are just higher level order errors. At least one simplification remains: We have not distinguished previous list *presentations* from previous list *recalls*. As Estes (1991) and Henson (1996) have noted, intrusions tend to come from previous recalls rather than presentations. If it is assumed that each recall is a new learning episode, however, little understanding would be gained and much complexity would be added to the model; instead we simply interpret the "list presentations" in the model as described as being produced by accurate recalls of previous lists.

Isolation and von Restorff Effects: Lippman (1980)

Many aspects of SIMPLE's behavior reflect the local distinctiveness of the locations that the episodic traces of items occupy in multidimensional psychological space. This general view predicts that if a single item within a list is made particularly distinctive along any dimension, then that item should be particularly memorable (cf. Figure 2C). This is the well-established isolation effect, or von Restorff phenomenon (for reviews, see Hunt, 1995; Wallace, 1965). Isolation effects have received a number of different interpretations, sometimes in terms of the establishment of "perceptual anchors" (e.g., Lippman, 1980) or perceptual salience and differential attention (see Hunt, 1995). SIMPLE offers a contrasting retrieval-based account.

We illustrate with data from Lippman (1980; Experiment 1). Lippman displayed a sequence of 12 consonant–vowel–consonant trigrams at a rate of 2 s per item. In the isolation condition, the seventh item was framed by a red rectangle. At test, participants were shown the 12 trigrams in random order and were required to estimate the ordinal position of each. The results, which are shown in Figure 21, were similar whether or not the seventh-presented trigram was enclosed by a red rectangle at test; Figure 21 shows only the conditions where the item was isolated at presentation but not retrieval (for similar results, see Bone & Goulet, 1968; Cimbalo, Nowak, & Soderstrom, 1981).

We addressed the data with SIMPLE. The temporal position of items' traces was set to the schedule of presentation; retrieval was assumed to commence after 0.1 s. Items were also represented along a log-transformed positional dimension, and the positional cue for the midseries item was given an increment of 10 in the condition where that item was distinctive. (An alternative approach could assume a third dimension in memory space; such an account behaves similarly.) The probability of item recall was calculated in the normal manner. The performance of the model is shown in Figure 21B. There were two free parameters, c and w_T.

The key assumption is that an isolated item is distinguished from other items in terms of its position along at least one dimension. The explanation is essentially the same as has been used to

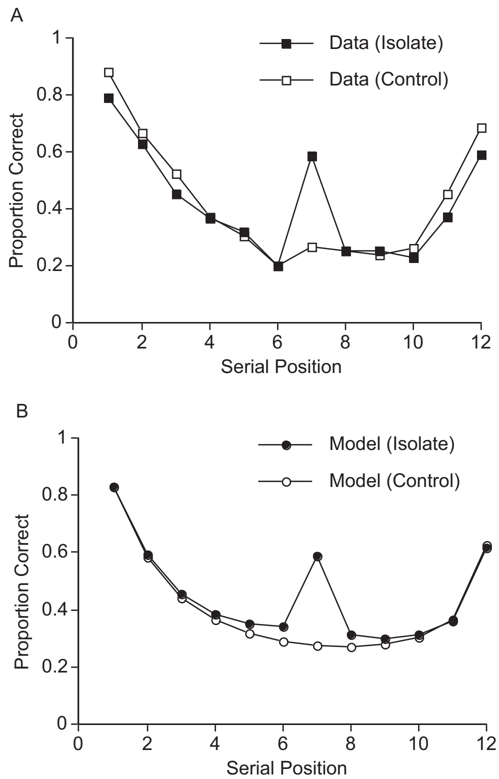

Figure 21. The isolation effect for recall of midlist items. A: Data adapted from Lippman (1980). B: Behavior of SIMPLE (scale-independent memory, perception, and learning model).

account for grouping effects, phonological similarity effects, and others. Thus, there is no need to assume any additional contrast-related attention or encoding devoted to the isolated item; the only difference in encoding of the isolated item relates to its different dimensional values, not its surprisingness per se (see Riefer & LaMay, 1998). Note that in the simulation above, SIMPLE typically predicts a slight advantage for the items adjacent to the isolated item. Recall of items neighboring a particularly isolated item is sometimes increased, and sometimes reduced (Wallace, 1965), with memory for subsequent items being impaired when the distinctive item is particularly attention demanding (Ellis, Detterman, Runcie, McCarver, & Craig, 1971; see Fabiani & Donchin, 1995, for discussion of the possible role of encoding in the von Restorff effect).

Grouping Effects: Hitch et al. (1996)

A key theoretical issue for time-based models of the type proposed here is their ability to accommodate hierarchical effects (Friedman, 2001). We have already addressed the model's ability to represent items in memory in terms of both their within-list position and their overall temporal distance from the point of recall. Other evidence for hierarchical representation in memory comes from grouping effects in memory for serial order (Frankish, 1985, 1989; Henson, 1998b; Hitch et al., 1996; Ng, 1996; Ng & Maybery, 2002; Ryan, 1969a, 1969b; Wickelgren, 1967). In a typical task, participants are presented with a series of nine items, either regularly or organized into subgroups of three items, and are then required to recall all of the items in correct serial order. Key findings are that (a) performance is higher overall when grouping occurs; (b) small primacy and recency effects are evident within each group, as well as at the level of the list as a whole; (c) grouping effects are larger when the structure is imposed through the insertion of temporal gaps between groups during list presentation (Ryan, 1969a); (d) auditory presentation gives rise to larger effects of grouping, but the effects discussed here are qualitatively similar for auditory and visual presentation (Frankish, 1985, 1989; Hitch et al., 1996; Ng, 1996); (e) the optimum group size is three (Wickelgren, 1967); and (f) many order errors preserve within-group position (Henson, 1998b; Ryan, 1969a, 1969b).

Many of these findings fall out reasonably naturally from the SIMPLE framework. The contrasting memory representations for the two-dimensional (grouped) case and the one-dimensional (ungrouped) case are assumed to be essentially similar to those in Figure 8, although with the vertical axis representing within-group position (grouped case only) and the horizontal axis representing temporal distance from recall. Note that the within-group position dimension is nontemporal. The figure illustrates the idea of a trade-off—grouping causes temporally adjacent items to become more distant from each other in the two-dimensional space and hence more memorable. Set against this, grouping can worsen performance insofar as items' memory representations may become closer to the representations of items that are not temporally adjacent but share the same within-group position. This reduction in performance can be evident in increased numbers of errors that preserve within-group position. Does SIMPLE capture the key effects?

We examined memory for a nine-item list in the model in both a grouped and an ungrouped condition. We followed the item presentation times adopted by Hitch et al. (1996). In the model of the grouped condition, there were three groups of three items. Onsets of items within a group were separated by 0.45 s, and each group was separated by 0.5 s. In the ungrouped condition, each item onset was separated by 0.6 s. With this schedule the total time to present the list was the same in both conditions (as in the experimental methodology of Hitch et al.). On the within-group position dimension, each item was given a value of 1, 2, or 3, corresponding to its within-group position. As omissions are possible, there were four free parameters: c; the attentional weight to the grouping dimension (i.e., $1 - w_T$), which was assumed to be greater (.22) in the case where the grouping dimension was highlighted by the temporal presentation schedule than in the ungrouped condition (.06); t; and s. Because the time course of output is not known, we made a similar assumption as in previous simulations, that is, that output time increased as an increasing function of output position: Specifically, the additional retrieval time for the nth item was assumed to be $n^{1.5}$ s.

The results are shown in Figure 22 for both grouped and ungrouped cases, along with the relevant conditions from Hitch et al. (1996). The basic pattern is reasonably similar to that observed experimentally when visual presentation of grouped and ungrouped lists is employed (Hitch et al., 1996): There is improved performance in the grouped condition overall, and there are within-group serial position effects. Performance did not fall off as fast with serial position in the model as in the data; this may reflect the operation of output interference (not incorporated in the present version of SIMPLE, but cf. Lewandowsky et al., 2004), progressively reducing encoding for successive items at presentation, or

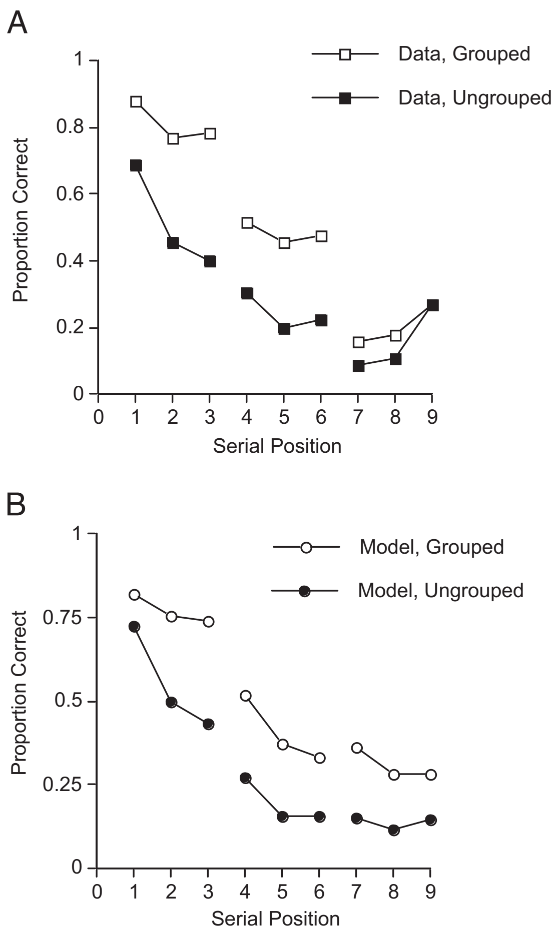

Figure 22. Grouping effects in immediate serial recall. A: Data adapted from Hitch et al. (1996). B: Behavior of SIMPLE (scale-independent memory, perception, and learning model).

may reflect PI from previous lists. We do not believe the data are yet sufficient to distinguish between these possibilities. The slowing of response latencies as recall progresses, which gives rise to the primacy that is observed, is in typical experiments probably less than we have assumed. We also note the relative lack of recency in the model.

The model's behavior with grouped lists can be understood in terms of principles already introduced. The overall decline in performance across serial positions, for both grouped and ungrouped lists, is due to the shifting temporal perspective of recall (more time has passed when the later items are recalled). The overall advantage for items in the grouped as opposed to the ungrouped list arises because of the extra dimension on which grouped items are represented. Thus, Items 3 and 4 (the last item of the first group and the first item of the second group) are close to one another, and hence not very distinctive, on the within-list temporal dimension. In the grouped case, in contrast, Items 3 and 4 are distant from one another on the dimension of within-group position and so become more distinctive and hence discriminable within memory. This also explains the mini primacy and recency effects that occur within each group, for the end-group items are more distinctive than are midgroup items on the position-within-group dimension. Representing items from the grouped list on the additional dimension of within-group position helps performance for the reasons given above, but also causes items that occupy the same within-group position (e.g., Items 4 and 7, or 2 and 5) to become closer together in psychological space. This should lead to a high proportion of order errors (exchanges) between items from the same within-group position, as is seen in the data whether presentation is auditory (Ryan, 1969a, 1969b) or visual (Henson, 1998b; Ng, 1996). We therefore examined the proportion of order errors produced by the model over different within-list separations, using the same parameters as above, and the results were as expected: Most errors (37%) were adjacent transpositions, but there were more movement errors that preserved within-group position (34%) than errors that involved movement of only two positions (24%). The precise numbers obtained vary substantially with parameter values, however.

General Discussion

We have outlined a temporal distinctiveness model of memory, motivated by the idea that memory retrieval involves temporal discrimination analogous to absolute identification, and argued that the core principles of the model provide a coherent perspective on a broad range of serial and free recall data. SIMPLE suggests that all forgetting is due to reduced local distinctiveness in psychological space and that no forgetting is due to trace decay. The same mechanisms are assumed to be used in retrieval from episodic memory as are used in absolute identification and categorization tasks. Moreover, and perhaps most important, the same mechanisms are suggested to govern retrieval over both short and long timescales for the data we have considered here and are claimed to underpin regularities in data derived across a range of timescales. We now briefly summarize (a) the key features and behaviors of the model, (b) its relation to other models, and (c) limitations and extensions.

Key Properties and Behavior

What are the general properties of SIMPLE that give rise to its qualitative behavior? Our aim has been to explain as much as possible with as few assumptions as possible, rather than to account for every nuance of the data. As applied to the simplest serial recall tasks, SIMPLE has just one free parameter: c. When stimuli are assumed to be represented in two dimensional space, a second attentional weight parameter, w_T, must be introduced. An additional attentional weighting parameter must be introduced for each new dimension. Two further parameters are needed when omissions are made possible by the nature of the experimental task (e.g., in free recall).

What is the explanatory value of these parameters? We are keenly aware of the danger of descending to mere curve fitting. Such concerns can be allayed in at least three ways. First, one can look for consistency in parameters across simulations. The c parameter is of particular interest, as it determines the rate at which confusability declines with temporal separation. Thus, the value of c should be related to the time span covered by the data being modeled. In modeling the near scale-similarity in absolute identification, Neath and Brown (2006) scaled the value of the c parameter in proportion to the range of stimuli to be identified. If, as hypothesized, the same account is relevant to scale-similar effects in memory retrieval, we would expect an association between estimated c values and the extent of the logarithmically compressed temporal dimension that is occupied by the to-be-remembered material. We therefore examined the best fit values of the key c parameter as a function of a proxy measure—the time span of recall (measured as the time from list start to commencement of retrieval)—for relevant simulations of serial and free recall in the present article.[12] We excluded parameter estimates from models that incorporated retrieval dynamics, used the shortest time span in experiments that included more than one, and used the estimate from the better model when two models were reported for the same data set.

The result is shown in Figure 23 ($R^2 = .90$ for the best fitting power-law relation illustrated). The effect is carried by the three extreme points corresponding to large c values; these estimates come from two separate studies (Nairne, 1992, and Underwood, 1957), but no clear relation is evident when the points are removed. Although this approach is very crude, ignoring as it does all methodological differences between experiments and the value of threshold and attentional weight parameters, it provides some suggestion that the c parameter may be related across different simulations to temporal features of the task environment.[13] On the other hand, it is also clear from Figure 23 that the value of the c parameter is far from completely predictable from the temporal parameters of a given experiment. It remains unclear how much this should be attributed to inadequacies in the model and how much to factors such as differences in extraexperimental interference between paradigms. Future research will be needed to test the model's assumption that (all other things being equal) different temporal list durations and retention intervals should be systematically related to c values.

Similar retrieval principles over different timescales. Development of SIMPLE was motivated by the general hypothesis that any comprehensive memory model will need to account for scale-similar effects as well as data previously taken to implicate different memory retrieval processes operating over different timescales. Thus, we have emphasized retrieval processes that operate over both short and long timescales.

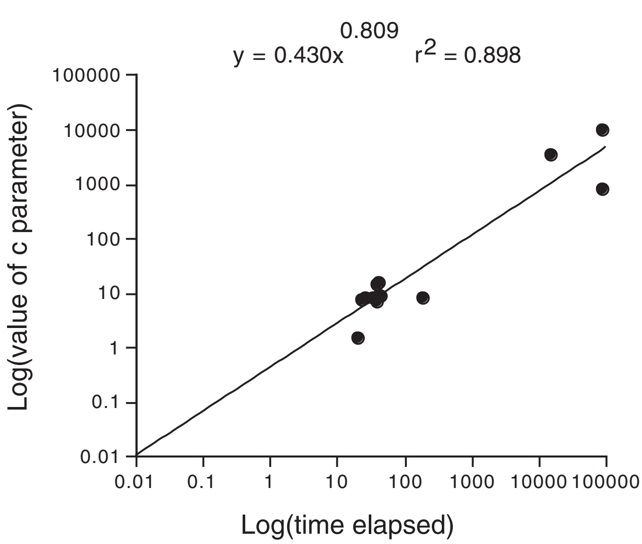

Figure 23. Estimates of the value of the c parameter as a function of experimental temporal range.

The claim of scale-similar memory amounts to the suggestion that the conventional distinction between short-term and long-term memory is not needed for the data we have considered. A detailed description of all of the findings that have been taken to support a short-term/long-term memory distinction (for comprehensive statements, see Atkinson & Shiffrin, 1968; Baddeley, 1976; Glanzer, 1972; Izawa, 1999) must be the subject of a separate article. We note in particular that arguments for separate stores are buttressed by many other methodologies that we have not considered in the present article (e.g., McElree, 1996; Wickelgren, Corbett, & Dosher, 1980).

However, we emphasize that we nowhere claim that all dimensions of psychological space are equally weighted at retrieval over all timescales. For example, due to Weberian compression the position of memory traces along a temporal distance dimension is likely to be particularly helpful in distinguishing those items from their near neighbors after relatively short effective retention intervals. Thus, as retention interval changes there may be a smooth and gradual shift in the attentional weights paid to different dimensions. Selective loss of the ability to represent items along some particular dimension can therefore lead to selective memory impairments without undermining the claim that similar interference-sensitive retrieval processes operate over a range of timescales, as can the inability to attend to a particular psychological dimension at retrieval (cf. Wickelgren, 1973).

More generally, many of the classic findings that have been taken to support different storage and retrieval principles over different timescales have recently been subject to reinterpretation (see, e.g., Laming, 2006, and Tan & Ward, 2000, for a unitary account of serial position curves and other free recall data). For example, neuropsychological data have long underpinned a conventional short-term/long-term memory distinction. However Brown, Della Sala, Foster, and Vousden (in press) have argued that the selective abolition of primacy in classic hippocampal amnesia can be understood in terms of a single-process temporal distinctiveness model when the pattern of rehearsal is taken into account (see also G. D. A. Brown & Lamberts, 2003).

A particularly relevant challenge to a unitary view comes from recent claims that long-term and short-term recency effects reflect the operation of different mechanisms (Davelaar et al., 2005, 2006). For example, Davelaar et al. (2005) enumerated several dissociations between long-term and short-term recency effects. Neath and Brown (2006), however, suggested that a number of these dissociations can in fact be accommodated (or indeed are predicted) by SIMPLE without abandonment of the principle that the same retrieval mechanisms operate over temporal frames normally associated with separate stores and/or forgetting mechanisms. Thus, the model we have presented does not claim to capture all data that have been taken to support a distinction

[12] An alternative and theoretically more appropriate method would relate c to the ratio of the largest and smallest temporal distances of to-be-remembered items; such analysis is not yet possible as the temporal dynamics of recall, which are crucial to the ratio calculation, are not known in many cases.

[13] This analysis suggests that c should be made to depend on list duration and/or retention interval when modeling data from a single experiment (e.g., variable list-length experiments) or examining the time course of forgetting.

between short- and long-term memory, but it does claim that several such data are susceptible to a unitary interpretation.

Interference-based forgetting. In contrast to many recent implemented and verbally described models of memory, SIMPLE assumes that there is no trace decay over time. The model can nevertheless explain the appearance of time-based forgetting as being due to increasing PI (the main source of retrieval failure in the model) over time.

The claim embodied in SIMPLE is straightforward. All forgetting, over both short and long timescales, is due to interference (see also Nairne, 2002). There is no trace decay in the model.[14] This absence of trace decay distinguishes SIMPLE sharply from the majority of recent implemented models of both short-term and working memory (e.g., Anderson, Bothell, Lebiere, & Matessa, 1998; Anderson & Matessa, 1997; Burgess & Hitch, 1992, 1999; Henson, 1998b; Kieras, Meyer, Mueller, & Seymour, 1999; Lovett, Reder, & Lebiere, 1999; Page & Norris, 1998; Schneider, 1999). Despite the absence of trace decay, SIMPLE can explain how forgetting may occur owing to the passage of time alone. This is because of the way recall perspective changes over time and is perhaps best understood in terms of Crowder's (1976) telephone pole model. As times passes, recall perspective on a list of items changes and the items appear less distinctive. Nothing in the items' representations has changed; they need not have decayed or degraded in any way for forgetting to occur.

Is the no-decay claim coherent? There are many different notions of decay. Most fall into one of two classes. The most common intuition is that "decay" must involve some change, over time, in the stored memory representations themselves. According to this type of definition, there is no trace decay in SIMPLE (although under extreme circumstances, such as head injury, physical disturbance of memory traces must always be possible at a physical level). However, a second class of definition focuses on decay in memory performance rather than decay of memory representations; according to such definitions, any forgetting that occurs owing to the passage of time alone qualifies as decay. For example, Peterson (1966) defined decay as "forgetting which would occur no matter how dissimilar preceding and intervening activities were to the tested material" (p. 199; see also Cowan et al., 2001). We hope that by showing how time-based forgetting can occur in the absence of trace decay or degradation of any kind, and by showing how the passage of time can lead to release from PI (simulations above; see also Estes, 1955; Mensink & Raaijmakers, 1988), SIMPLE illustrates how key features of the data can be explained without the assumption of trace decay, spontaneous recovery of associations, consolidation, or a "Factor X" (Melton & Irwin, 1940).

Relation between different memory tasks. It has been suggested that similar retrieval mechanisms operate during memory tasks that have previously most often been treated as distinct. In particular, it is suggested that the contrasting serial position curves in forward serial recall, free recall, and probed serial recall reflect the differing temporal requirements of recall across the tasks. Broadly speaking, SIMPLE predicts that both primacy and recency will be observed in both free and serial recall tasks (owing to edge effects). Edge effects will generally be greater for shorter lists, but they may be reduced when performance is low and many omissions occur or enhanced (particularly in the case of primacy) through rehearsal. More interesting, the Weberian compression of the temporal dimension predicts much greater recency than primacy whenever the most recently presented items can be retrieved soon after their presentation (probed serial recall; immediate free recall; immediate recognition) but, conversely, predicts primacy to be greater than or equal to recency when late-presented items must have their recall postponed, thus causing output interference or the passage of time and PI to impair retrieval of late-presented items (delayed free recall; forward serial recall). The overall pattern of data appears broadly compatible with these conclusions, which should apply to memory for items of all types.

The form of the forgetting function. No simple single equation governs the form of SIMPLE's forgetting function. Despite expenditure of a considerable amount of ingenuity and empirical effort, the data appear at least consistent with the same conclusion (see, e.g., Chechile, 2006; Rubin et al., 1999; Rubin & Wenzel, 1996; T. D. Wickens, 1999). The forgetting curve in SIMPLE is closely approximated by exponential forgetting in the short term and power-law forgetting over longer time periods, but the form of the best fitting function was found to depend to a large (and perhaps intuitively surprising) extent on parameter values that, from a theoretical point of view, seem rather peripheral to the core assumptions of the model. We therefore suggest that the search for "the" forgetting function may be misguided.

Multiple traces in memory. A key theoretical claim of SIMPLE is that a separate trace is stored in memory for each episode of item occurrence (cf. Hintzman, 1976, 1986). This distinguishes SIMPLE from several other recent models (e.g., Farrell & Lewandowsky, 2002) while aligning it with numerous previous exemplar models of memory. Indeed, according to SIMPLE it is precisely location along the temporal dimension of memory that keeps traces of items apart. There may be good adaptive reasons for preserving distinct traces of multiple episodes, for the counting of such traces is often assumed to be important in estimation and calculation (e.g., Gallistel, 1990; Gigerenzer, 2000). Many of the difficulties of "global" models of memory (Clark & Gronlund, 1996), whether applied to recognition, serial recall, or free recall, appear to result from the agglomeration of separate episodes into a single memory or weight matrix. Crowder (1976) reviewed much relevant evidence; particularly decisive points in the present context include the fact that participants can reliably distinguish separate repetitions of an item and locate each occurrence separately (Hintzman & Block, 1971) and the fact that recency and frequency can generally although not always be distinguished (e.g., Flexser & Bower, 1974; Morton, 1968). Note that the "separate traces" issue is separate from the question of whether distinctive temporal–contextual tagging information is stored in memory—it is possible for item-to-context associations to be stored together in a single memory matrix (e.g., G. D. A. Brown et al., 2000) or for the same associations to be formed but stored separately (e.g., G. D. A. Brown, Vousden, McCormack, & Hulme, 1999; Vousden, Brown, & Harley, 2000). The arguments motivating SIMPLE argue for the latter possibility.

Relation to Other Models

The resulting model has affinities with a number of previous theoretical approaches. It can be viewed variously as (a) an exten-

[14] We do not exclude the possibility that a formally equivalent decay-based formulation could be given.

sion of Murdock's (1960) distinctiveness theory that accommodates time-based and local neighborhood effects; (b) a generalization and extension of early temporal discriminability and ratio models of memory (e.g., Baddeley, 1976; Bjork & Whitten, 1974; Crowder, 1976; Glenberg & Swanson, 1986); (c) the addition of a temporal dimension to a simplified exemplar model (Nosofsky, 1986, 1992); (d) the extension into the domain of temporal memory of the feature model's use of the Luce choice model for cue-driven recall (Nairne, 1990); (e) an extension of Neath's (1993a, 1993b) temporal distinctiveness model to allow isolation effects and primacy effects to be accounted for; or (f) a more analytic and abstract version of recent oscillator-based and contextual overlap models of memory for temporal order (e.g., G. D. A. Brown et al., 2000) and speech production (Vousden et al., 2000); see Howard and Kahana (2002) for an alternative temporal–contextual approach to free recall. More generally, following Gallistel (1990), the model places time and temporal interference at the center of memory and relates memory retrieval to perceptual discriminability. Indeed, there are strong resonances between the Gallistel and Gibbon (2002) nonassociative model of animal learning and the current nonassociative model of human memory.

In particular, SIMPLE imports the explanatory mechanisms previously developed to account for long-term memory tasks such as categorization, as well as recognition memory and absolute identification, to the domain of serial and free recall. Current models of absolute identification, categorization, and recognition performance account for a range of empirical data to a high level of precision (e.g., Ashby, 1992; Ashby & Perrin, 1988; Erickson & Kruschke, 1998; Estes, 1994; Kruschke, 1992; Kruschke & Johansen, 1999; Lamberts, 1995; Nosofsky, 1986; Nosofsky & Palmeri, 1997). In many respects such models seem more advanced than current models of serial and free recall.[15] However, the insights embodied in models of identification and classification have not generally been applied to traditional serial and free recall memory paradigms. Here we suggest that this is partly because multidimensional scaling models of categorization have not included *time* as an important dimension underpinning memory retrieval. Although models can allow for the differential availability in memory of exemplars,[16] the relation between temporal factors and memory/exemplar availability has not been widely explored.

The model clearly has close affinities with ratio-rule models of memory (e.g., Baddeley, 1976; Bjork & Whitten, 1974; Crowder, 1976; Glenberg et al., 1983; Koffka, 1935) in that temporal ratios determine the discriminability of items. However, SIMPLE makes different predictions from these prior models, primarily because it assumes that several near temporal neighbors, rather than a single preceding item in the list, determine retrieval difficulty. Neath and Brown (2007) provide extensive discussion of the relation between SIMPLE and ratio-rule models.

More generally, SIMPLE is naturally viewed as a temporal distinctiveness model of memory. The concept of distinctiveness is used in many different senses in the memory literature and stands in need of rigorous definition if circularity ("better recalled memories are more distinctive") is to be avoided. Murdock (1960) introduced just such a definition; the present model offers an alternative. In intuitive terms, SIMPLE states that "distinctive" memories will be those that occupy relatively isolated regions of psychological space. The notion that "crowded" materials will be remembered less well than "isolated" materials has a long history (e.g., Buxton & Newman, 1940) and has often, but not always, been taken to support some form of intraserial interference similar to the type explored in this article (McGeoch & Irion, 1952). SIMPLE provides a formalization of the notion of distinctiveness that derives from categorization theory.

How does SIMPLE relate to specific extant models of short-term memory (Burgess, 1995; Burgess & Hitch, 1996, 1999; Henson, 1998b; Henson et al., 1996; Houghton, 1990, 1994; Page & Norris, 1998)? There are clear points of contrast between SIMPLE and almost all previous models. First, SIMPLE makes the claim that the same retrieval principles apply over all timescales. Second, as already noted, SIMPLE claims that no forgetting due to trace decay occurs. Third, SIMPLE eschews the assumption of reducing attention, activation, or encoding to explain primacy effects. Fourth, in contrast to the model of Lewandowsky and Murdock (1989), no item–item associations are assumed. In what follows we focus on a small number of models of serial recall, as these provide the clearest contrasts on some of the key principles, but we acknowledge the large number of models of memory that we cannot do justice to here for reasons of space (e.g., Anderson et al., 1998; Gillund & Shiffrin, 1984; Howard & Kahana, 2002).

In its emphasis on the importance of a temporal dimension in serial recall, SIMPLE bears a family resemblance to the OSCAR model (G. D. A. Brown & Vousden, 1998; G. D. A. Brown et al., 1999, 2000; Maylor, Vousden, & Brown, 1999; Vousden & Brown, 1998; Vousden et al., 2000) and the Burgess and Hitch model (Burgess & Hitch, 1996, 1999; see also Burgess & Hitch, 1992, and Howard & Kahana, 2002, for time and context in free recall). These models assume that hierarchical contextual signals may underpin short-term serial recall. The models differ in terms of their psychological interpretation and the data they address: The Burgess and Hitch model is intended as a model of the phonological loop component of working memory (e.g., Baddeley, 1986), whereas OSCAR is intended to apply to serial recall tasks over all timescales but, unlike the Burgess and Hitch model, does not directly address the data motivating the phonological loop account. SIMPLE differs from both models in its level of abstraction; it also differs from the Burgess and Hitch model in its denial of trace decay and in its emphasis on a logarithmically transformed time dimension. OSCAR can in many respects be seen as a mechanism-level instantiation of SIMPLE, as noted above; indeed the development of SIMPLE was motivated by the desire to capture the key explanatory elements of OSCAR within a simpler and more tractable framework. SIMPLE differs from earlier time-tagging models (Hintzman & Block, 1971) in its emphasis on temporal distance from the point of retrieval. One other model shares SIMPLE's emphasis on the role of time in memory: the temporal distinctiveness model of Neath and his colleagues (Neath, 1993a, 1993b). SIMPLE differs from the Neath temporal distinctiveness model in its assumption that local, rather than global, temporal distinctiveness is what governs performance. SIMPLE contrasts with the primacy gradient model (Page & Norris, 1998) particularly in its

[15] For example, such models are better able to accommodate issues like the dangers of averaging data over participants (Ashby, Maddox, & Lee, 1994; Maddox, 1999).

[16] For example, they can do so via the *M* parameter in the Nosofsky and Palmeri (1997) exemplar-based random walk model.

emphasis on commonalities between short-term and long-term memory and in its specification of the role of time.

In the SIMPLE model, there is no mutual exclusion between temporal and positional representations. SIMPLE assumes that the concept of distinctiveness along a temporal dimension is essential to the explanation of many key phenomena, and differs in this from recent purely positional models (e.g., Henson et al., 1996; Henson, 1998b). However, as noted, the multidimensional psychological space assumed by SIMPLE can naturally be extended to enable representation of items along a positional as well as a temporal dimension; although temporal organization is assumed to be primary, we assume that psychological space will become organized along whatever dimensions are most accessible and useful for a particular task at hand. In particular, the inclusion of a nontemporal positional dimension may be required (Henson, 1999; Lewandowsky et al., 2004; Ng & Maybery, 2002).

Finally, we note that the more general potential to attach different weights to different retrieval dimensions (e.g., temporal vs. nontemporal) may allow a multidimensional model to behave more or less episodically depending on the relative weight given to the episodic/temporal dimension (see G. D. A. Brown & McCormack, 2006; Humphreys, Bain, & Pike, 1989).

Perturbation model. The perturbation model (PM) developed by Estes (1972, 1985, 1997) emphasizes the distortion in memories' attribute values over time or over retrieval attempts. Descriptions of the model typically focus on the perturbations of items' codes on a positional, rather than temporal, dimension, but this focus need not be seen as an intrinsic feature of the model. The perturbation model is similar to SIMPLE in the way it treats the temporal/positional location of an item as an attribute that can be remembered or forgotten like any other. However, the perturbation model differs from SIMPLE in its emphasis on processes (perturbations) applying to memory representations. The perturbation model assumes that stored memories become distorted; SIMPLE in contrast emphasizes retrieval-stage interference. A further difference concerns the relation between forgetting and the passage of time. The amount of perturbation of memories' dimensional attributes is not assumed by the perturbation model to be a function of the passage of time alone; rather, successive retrieval attempts will increase the probability or number of perturbations (Estes, 1997). In SIMPLE, in contrast, forgetting may occur (all other things being equal) whether or not successive retrievals have intervened. SIMPLE does not incorporate the dual-trace assumptions that form an important part of the most recent statement of the perturbation model and does not at present address the data that the dual-trace assumption was intended to explain. Perhaps the most fundamental point of contrast between the perturbation model and SIMPLE, however, concerns the theoretical treatment of the *relation* between time and other dimensions along which items are represented. In the perturbation model, perturbations are seen as additional to and separate from the multidimensional featural representations of objects in the array model of recognition and categorization (Estes, 1994). In other words, in the perturbation model there are two assumptions: (a) Items are represented as vectors of features in multidimensional space, and (b) featural values may perturb over time. In SIMPLE, in contrast, time is treated as just like any other dimension; an item's distinctiveness along a temporal distance dimension (given a particular recall perspective) affects recall in exactly the same manner as does the item's distinctiveness along any other dimension. Despite these large differences in the ways the perturbation model and SIMPLE are interpreted psychologically, they often make very similar predictions in practice if certain of SIMPLE's assumptions are incorporated into the perturbation model.

Feature model. Central to the feature model is a distinction between modality-dependent and modality-independent features (Nairne, 1988, 1990; Nairne et al., 1997; Neath, 2000; Neath & Nairne, 1995). This reflects the feature model's initial focus on accounting for modality effects, which have not been examined in SIMPLE. SIMPLE and the feature model share the assumption that items are located in multidimensional space, although the feature model uses binary features (akin to the model of Medin & Schaffer, 1978; see also Estes, 1994) whereas SIMPLE assumes continuous-valued dimensions (akin to the model of Nosofsky, 1986). Both models assume that the effectiveness of retrieval cues will depend on the extent to which they cue a given memory relative to the extent to which they cue other, competing memories. The models differ in the assumed source of forgetting: In the feature model, primary memory is conceived of as a repository for retrieval cues, and forgetting occurs because of overwriting of cues rather than, as in SIMPLE, Weberian compression. However, SIMPLE and the feature model share the important assumption that no trace decay need be assumed.

Limitations and Extensions

A full list of limitations would be extensive; we focus here on the additional mechanisms that would be needed to account for phenomena closest to those to which SIMPLE has already been applied. We believe that a strength of the SIMPLE model is the possibility it offers of a rapprochement between models of categorization and models of episodic memory. However, the modeling framework we have adopted is in many respects rather simpler than is typically used in modern categorization models. In particular, we have not explored the issue of response determinism at length, nor have we calculated psychological distances other than by a simple city block metric. These simplifications have been adopted largely because we were able to account for the key qualitative phenomena quite adequately without further parameters. However, Surprenant et al. (2006) derived multidimensional scaling solutions based on the memory confusions when younger and older subjects recalled lists of acoustically confusable and nonconfusable items. SIMPLE was shown to account for both the overall difference in performance between the two age groups and the difference between acoustically confusable and nonconfusable items largely in terms of the multidimensional scaling coordinates.

We also note that SIMPLE contains no mechanism for varying the strength with which items are encoded as a function of serial position. Although, for the serial position data that we have considered, we have not found it necessary to assume reduced attention or encoding for successive items in a list (in contrast to other models of the same data), both intuition and empirical considerations suggest that there may be a role for such parameters in extending the model. For example, several models of human and animal learning incorporate the intuition that greater encoding will occur for items that are somehow surprising or unexpected in a given context, and the tendency for retrievability of primacy items to increase in absolute terms after a delay (see Bjork, 2001, for a review) may point to the need for inhibitory or encoding mechanisms not yet incorporated into SIMPLE. Similar encoding-level considerations arise in the context of explanations of distributed and massed practice effects (e.g., K. Braun & Rubin, 1998).

A further issue concerns learning and practice. In the present article, we have applied SIMPLE almost exclusively to cases of single-trial learning, where one presentation of a list of nonrepeated items is followed by recall. An area for future research must involve extension of the model to cases involving multiple presentations and associated learning and transfer effects. Initial exploration through simulation suggests that the retrieval assumptions of SIMPLE may combine well with assumptions of multiple-trace models such as that of Logan (1988) or Anderson, Fincham, and Douglass (1999) and that a position-from-start dimension may be important in accounting for cross-list transfer effects (see, e.g., Chen, Swartz, & Terrace, 1997; Hitch & Fastame, 2005). We postpone detailed consideration.

Response suppression mechanisms, and their role in explaining errors in recall in short-term memory paradigms, are now quite well understood (e.g., Henson, 1998a; Lewandowsky, 1999; Vousden & Brown, 1998). Little real explanatory gain would be achieved by incorporating such mechanisms into SIMPLE, although there is in principle no difficulty in doing so. Indeed, we have implemented a stochastic version of the model, with response suppression included, and the essential behavior of the model is the same. Additional mechanisms are needed to explain modality effects (see Penney, 1989, for a review), and Bayesian redintegration processes along with richer multidimensional semantic representations would need to be combined with the temporal dimension described in the present article to provide a complete account of lexicality and frequency effects in short-term memory in terms of local distinctiveness (Hulme, Maughan, & Brown, 1991).

In its application to free recall, and perhaps also serial recall, augmentation of SIMPLE will be needed to accommodate the full range of effects of rehearsal on the form of the serial position function and, in particular, the preservation of primacy after a filled retention interval (Tan & Ward, 2000). A noteworthy simplification in the model is the treatment of items' temporal locations as point sources; a complete account will need to reflect items' temporal extension. Along related lines, cumulative rehearsal of items in the model to produce multiple traces can have the paradoxical consequence of making each individual trace less isolated and hence less retrievable; the summed temporal extension of an item's rehearsals provides a more complete (albeit more complex) account (G. D. A. Brown & Morin, 2006). The present model simplifies considerably by ignoring such considerations and can perhaps best be viewed as a model of rehearsal-free memory performance. The aim has been to understand the operation and consequences of temporal distinctiveness principles rather than to specify a full process model.

A further limitation of the model as described is its silence on organizational factors in free recall (e.g., Bousfield, 1953; Mandler, Pearlstone, & Koopmans, 1969) and the emergence of such factors with practice (Tulving, 1966). This limitation arises in part because we have omitted any specification of output order processes in SIMPLE. However, we note that the model treats temporal and semantic dimensions in the same way; items can be similar in their location along the temporal dimension in just the same way that they can be similar in their location along a semantic dimension or dimensions. Therefore, any tendency to recall similar items in sequential clusters (e.g., if the location in memory cued during recall of item $n + 1$ tends to be close to the location in memory cued during recall of item n) will lead to clustering along both semantic and temporal dimensions. Just this pattern is seen in the data: Not only do semantically related items tend to be recalled together, but items that were temporally close at presentation also tend to be recalled together (e.g., Howard & Kahana, 1999; Kahana, 1996). In SIMPLE, according to which the temporal dimension behaves like any other, both types of clustering in recall—semantic and temporal—can arise from the same mechanism.

We have already noted applications of SIMPLE to amnesia data (G. D. A. Brown, Della Sala, et al., in press; G. D. A. Brown & Lamberts, 2003) and dissociations between short- and long-term recency effects (Neath & Brown, 2006). The consequences of emphasizing a positional rather than (or in addition to) a temporal dimension in serial recall, along with the role of output interference, have not been considered here but are dealt with at length in Lewandowsky et al. (2004, 2006). In addition to SIMPLE's application to isolation effects in absolute identification (Neath et al., 2006, Neath and Brown (2006) apply the model to shifts from recency to primacy as a function of retention interval (e.g., Neath & Knoedler, 1994), differing schedules of presentation in serial recall (Neath & Crowder, 1990), additional data concerning order memory from milliseconds to weeks (Huttenlocher et al., 1992) and serial position effects in semantic memory (the order of verses in popular hymns: Maylor, 2002), and the interaction of phonological confusability effects with delay (Nairne & Kelley, 1999).

Conclusion

In summary, the model we have developed takes as its starting point the ideas that (a) many important human memory phenomena appear similar over a wide range of timescales, and (b) a useful starting point in accounting for such phenomena may be the assumption, contrasting with assumptions that are made in most existing literature, that the same principles may govern forgetting and retrieval over both short and long timescales. We also assumed that memory retrieval requires discrimination akin to that observed in absolute identification, and hence that serial position curves in identification and recall tasks have a common origin. The resulting model uses principles independently derived in models of long-term memory tasks such as categorization to account for a range of data from both serial and free recall tasks.

References

Alin, L. H. (1968). Proactive inhibition as a function of the time interval between the learning of the two tasks and the number of prior lists. *Journal of Verbal Learning and Verbal Behavior, 7,* 1024–1029.

Alluisi, E. A., & Sidorsky, R. C. (1958). The empirical validity of equal discriminability scaling. *Journal of Experimental Psychology, 55,* 86–95.

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38,* 341–380.

Anderson, J. R., Fincham, J. M., & Douglass, S. (1999). Practice and retention: A unifying analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1120–1136.

Anderson, J. R., & Matessa, M. (1997). A production system theory of serial memory. *Psychological Review, 104,* 728–748.

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review, 96,* 703–719.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2,* 396–408.

Ashby, F. G. (1992). Multidimensional models of categorization. In F. G. Ashby (Ed.), *Multidimensional models of perception and cognition* (pp. 449–483). Hillsdale, NJ: Erlbaum.

Ashby, F. G., & Maddox, W. T. (1993). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology, 37,* 372–400.

Ashby, F. G., Maddox, W. T., & Lee, W. W. (1994). On the dangers of averaging across subjects when using multidimensional-scaling or the similarity-choice model. *Psychological Science, 5,* 144–151.

Ashby, F. G., & Perrin, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychological Review, 6,* 363–378.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 2, pp. 89–195). New York: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1971, August). The control of short-term memory. *Scientific American, 224,* 82–90.

Baddeley, A. D. (1968). How does acoustic similarity influence short-term memory? *Quarterly Journal of Experimental Psychology, 20,* 249–263.

Baddeley, A. D. (1976). *The psychology of memory.* New York: Basic Books.

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

Baddeley, A. D., & Ecob, J. R. (1970, July 18). Simultaneous acoustic and semantic coding in short-term memory. *Nature, 227,* 288–289.

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

Baddeley, A. D., & Scott, D. (1971). Short term forgetting in the absence of proactive interference. *Quarterly Journal of Experimental Psychology, 23,* 275–283.

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

Barenblatt, G. I. (1996). *Scaling, self-similarity, and intermediate asymptotics.* Cambridge, England: Cambridge University Press.

Bjork, R. A. (2001). Recency and recovery in human memory. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 211–232). Washington DC: American Psychological Association.

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

Bone, R. N., & Goulet, L. R. (1968). Serial position and the Von Restorff isolation effect. *Journal of Experimental Psychology, 76,* 494–496.

Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology, 49,* 229–240.

Bower, G. H. (1971). Adaptation-level coding of stimuli and serial position effects. In M. H. Appley (Ed.), *Adaptation-level theory* (pp. 175–201). New York: Academic Press.

Braun, H. W., & Heymann, S. P. (1958). Meaningfulness of material, distribution of practice, and serial-position curves. *Journal of Experimental Psychology, 56,* 146–150.

Braun, K., & Rubin, D. C. (1998). The spacing effect depends on an encoding deficit, retrieval, and time in working memory: Evidence from once-presented words. *Memory, 6,* 37–65.

Brodie, D. A., & Prytulak, L. S. (1975). Free recall curves: Nothing but rehearsing some items more or recalling them sooner? *Journal of Verbal Learning and Verbal Behavior, 14,* 549–563.

Brown, G. D. A., & Chater, N. (2001). The chronological organisation 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, England: Oxford University Press.

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

Brown, G. D. A., & Hulme, C. (1995). Modeling item length effects in memory span: No rehearsal needed? *Journal of Memory and Language, 34,* 594–621.

Brown, G. D. A., & Lamberts, K. (2003). Double dissociations, models, and serial position curves. *Cortex, 39,* 148–152.

Brown, G. D. A., & McCormack, T. (2006). The role of time in human memory and binding: A review of the evidence. In H. D. Zimmer, A. Mecklinger, & U. Lindenberger (Eds.), *Binding in human memory: A neurocognitive approach* (pp. 251–290). Oxford, England: Oxford University Press.

Brown, G. D. A., McCormack, T., Smith, M., & Stewart, N. (2005). Identification and bisection of temporal durations and tone frequencies: Common models for temporal and non-temporal stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 31,* 919–938.

Brown, G. D. A., & Morin, C. (2006, July). *Memory by temporal sampling.* Paper presented at the 4th International Conference on Memory, Sydney, Australia.

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

Brown, G. D. A., & Vousden, J. (1998). Adaptive sequential behaviour: Oscillators as rational mechanisms. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 165–193). Oxford, England: Oxford University Press.

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.

Brown, J. (1958). Some test of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10,* 12–21.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience, 6,* 755–765.

Burgess, N. (1995). A solvable connectionist model of immediate recall of ordered lists. *Advances in Neural Information Processing Systems, 7,* 51–58.

Burgess, N., & Hitch, G. J. (1992). Towards a network model of the articulatory loop. *Journal of Memory and Language, 31,* 429–460.

Burgess, N., & Hitch, G. J. (1996). A connectionist model of STM for serial order. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 51–72). Hove, England: Psychology Press.

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.

Buxton, C. E., & Newman, E. B. (1940). The forgetting of "crowded" and "isolated" materials. *Journal of Experimental Psychology, 26,* 180–198.

Chater, N., & Brown, G. D. A. (1999). Scale invariance as a unifying psychological principle. *Cognition, 69,* B17–B24.

Chechile, R. A. (2006). Memory hazard functions: A vehicle for theory development and test. *Psychological Review, 113,* 31–56.

Chen, S. F., Swartz, K. B., & Terrace, H. S. (1997). Knowledge of the ordinal position of list items in rhesus monkeys. *Psychological Science, 8,* 80–86.

Cimbalo, R. S., Nowak, B. I., & Soderstrom, J. A. (1981). The isolation effect in children's short term memory. *Journal of General Psychology, 105,* 215–223.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3,* 37–60.

Conrad, R. (1960). Serial order intrusions in immediate memory. *British Journal of Psychology, 51,* 45–48.

Conrad, R. (1964). Acoustic confusions in immediate memory. *British Journal of Psychology, 55,* 75–84.

Conrad, R. (1967). Interference or decay over short retention intervals. *Journal of Verbal Learning and Verbal Behavior, 6,* 49–54.

Cowan, N., Saults, J. S., & Nugent, L. D. (1997). The role of absolute and relative amounts of time in forgetting within immediate memory: The case of tone pitch comparisons. *Psychonomic Bulletin & Review, 4,* 393–397.

Cowan, N., Saults, J. S., & Nugent, L. (2001). The ravages of absolute and relative amounts of time on memory. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 315–330). Washington DC: American Psychological Association.

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

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

Davelaar, E. J., Haarmann, H. J., Goshen-Gottstein, Y., & Usher, M. (2006). Semantic similarity dissociates short- from long-term recency effects: Testing a neurocomputational model of list memory. *Memory & Cognition, 34,* 323–334.

DeSoto, C. B., & Bosley, J. G. (1962). The cognitive structure of a social structure. *Journal of Abnormal and Social Psychology, 64,* 303–307.

Ebenholtz, S. M. (1963). Position mediated transfer between serial learning and a spatial discrimination task. *Journal of Experimental Psychology, 65,* 603–608.

Ellis, N. R., Detterman, D. K., Runcie, D., McCarver, R. B., & Craig, E. M. (1971). Amnesic effects in short-term memory. *Journal of Experimental Psychology, 89,* 357–361.

Elvevåg, B., Brown, G. D. A., McCormack, T., Vousden, J. I., & Goldberg, T. E. (2004). Identification of tone duration, line length and letter position: An experimental approach to timing and working memory deficits in schizophrenia. *Journal of Abnormal Psychology, 113,* 509–521.

Erickson, M. A., & Kruschke, J. K. (1998). Rules and exemplars in category learning. *Journal of Experimental Psychology: General, 127,* 107–140.

Eriksen, C. W., & Hake, H. W. (1955). Absolute judgments as a function of stimulus range and number of stimulus and response categories. *Journal of Experimental Psychology, 49,* 323–332.

Eriksen, C. W., & Hake, H. W. (1957). Anchor effects in absolute judgments. *Journal of Experimental Psychology, 53,* 132–138.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psychological Review, 62,* 145–154.

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161–190). Washington, DC: Winston.

Estes, W. K. (1973). Phonemic coding and rehearsal in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 12,* 360–372.

Estes, W. K. (1985). Memory for temporal information. In J. A. Michon & J. L. Jackson (Eds.), *Time, mind, and behavior* (pp. 151–168). Berlin, Germany: Springer-Verlag.

Estes, W. K. (1991). On types of item coding and source of recall in short-term memory. In W. E. Hockley & S. Lewandowsky (Eds.), *Relating theory and data: Essays on human memory in honor of Bennet B. Murdock* (pp. 155–174). Hillsdale, NJ: Erlbaum.

Estes, W. K. (1994). *Classification and cognition.* New York: Oxford University Press.

Estes, W. K. (1997). Processes of memory loss, recovery, and distortion. *Psychological Review, 104,* 148–169.

Fabiani, M., & Donchin, E. (1995). Encoding processes and memory organization: A model of the von Restorff effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 224–240.

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

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

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

Flexser, A. J., & Bower, G. H. (1974). How frequency affects recency judgments: A model for recency discrimination. *Journal of Experimental Psychology, 103,* 706–716.

Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language, 24,* 200–209.

Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 469–479.

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, England: Oxford University Press.

Gallistel, C. (1990). *The organization of learning.* Cambridge, MA: MIT Press.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107,* 289–344.

Gallistel, C. R., & Gibbon, J. (2002). *The symbolic foundations of conditioned behavior.* Mahwah, NJ: Erlbaum.

Gardiner, J. M., Craik, F. I. M., & Birtwistle, J. (1972). Retrieval cues and release from proactive inhibition. *Journal of Verbal Learning and Verbal Behavior, 11,* 778–783.

Garner, W. G. (1962). *Uncertainty and structure as psychological concepts.* London: Wiley.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. G. Allan (Eds.), *Timing and time perception* (pp. 52–77). New York: New York Academy of Sciences.

Gigerenzer, G. (2000). *Adaptive thinking: Rationality in the real world.* Oxford, England: Oxford University Press.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review, 91,* 1–67.

Glanzer, M. (1972). Storage mechanisms in recall. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 5, pp. 129–193). New York: Academic Press.

Glanzer, M., & Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of Verbal Learning and Verbal Behavior, 5,* 351–360.

Glanzer, M., & Meinzer, A. (1967). The effects of intralist activity on free recall. *Journal of Verbal Learning and Verbal Behavior, 6,* 928–935.

Glenberg, A. M., Bradley, M. M., Kraus, T. A., & Renzaglia, G. J. (1983). Studies of the long-term recency effect: Support for a contextually guided retrieval hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9,* 231–255.

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

Glenberg, A. M., & Swanson, N. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 3–24.

Hacker, M. J. (1980). Speed and accuracy of recency judgments for events in short-term memory. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 651–675.

Harper, D. N., McLean, A. P., & Dalrymple-Alford, J. C. (1993). List item memory in rats: Effects of delay and task. *Journal of Experimental Psychology: Animal Behavior Processes, 19,* 307–316.

Haubensak, G. (1992). The consistency model: A process model for absolute judgments. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 303–309.

Healy, A. F. (1974). Separating item from order information in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 644–655.

Healy, A. F. (1975). Coding of temporal–spatial patterns in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14,* 480–495.

Healy, A. F., Havas, D. A., & Parker, J. T. (2000). Comparing serial position effects in semantic and episodic memory using reconstruction or order tasks. *Journal of Memory and Language, 42,* 147–167.

Healy, A. F., & Parker, J. T. (2001). Serial position effects in semantic memory: Reconstructing the order of the U.S. presidents and vice presidents. In H. L. Roediger III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 171–188). Washington DC: American Psychological Association.

Henson, R. N. A. (1996). *Short-term memory for serial order.* Unpublished doctoral dissertation, University of Cambridge, Cambridge, England.

Henson, R. N. A. (1998a). Item repetition in short-term memory: Ranschburg repeated. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1162–1181.

Henson, R. N. A. (1998b). 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.

Henson, R. N. A., Norris, D. G., Page, M. P. A., & Baddeley, A. D. (1996). Unchained memory: Error patterns rule out chaining models of immediate serial recall. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 49*(A), 80–115.

Hintzman, D. L. (1976). Repetition and memory. *Psychology of Learning and Motivation, 10,* 47–91.

Hintzman, D. L. (1986). "Schema abstraction" in a model-trace memory model. *Psychological Review, 93,* 411–428.

Hintzman, D. L., & Block, R. A. (1971). Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology, 88,* 297–306.

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology,* 49(A), 116–139.

Hitch, G. J., & Fastame, M. C. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory, 13,* 247–258.

Hitch, G. J., Rejman, M. H., & Turner, N. C. (1980, July). *A new perspective on the recency effect.* Paper presented at the meeting of the Experimental Psychology Society, Cambridge, England.

Houghton, G. (1990). The problem of serial order: A neural network model of sequence learning and recall. In R. Dale, C. Mellish, & M. Zock (Eds.), *Current research in natural language generation* (pp. 287–319). London: Academic Press.

Houghton, G. (1994). Inhibitory control of neurodynamics: Opponent mechanisms in sequencing and selective attention. In M. Oaksford & G. D. A. Brown (Eds.), *Neurodynamics and psychology* (pp. 107–155). London: Academic Press.

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 923–941.

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

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for words and non-words: Evidence for a long-term memory contribution to short-term memory tasks. *Journal of Memory and Language, 30,* 685–701.

Humphreys, M. S., Bain, J. D., & Pike, R. (1989). Different ways to cue a coherent memory system: A theory for episodic, semantic, and procedural tasks. *Psychological Review, 96,* 208–233.

Hunt, R. R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review, 2,* 105–112.

Huttenlocher, J., Hedges, L. V., & Prohaska, V. (1992). Memory for day of the week: A 5 + 2 day cycle. *Journal of Experimental Psychology: General, 121,* 313–326.

Izawa, C. (Ed.). (1999). *On human memory: Evolution, progress, and reflections on the 30th anniversary of the Atkinson–Shiffrin model.* Mahwah, NJ: Erlbaum.

Jensen, A. R. (1962). Temporal and spatial effects of serial position. *American Journal of Psychology, 75,* 390–400.

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

Kent, C., & Lamberts, K. (2005). An exemplar account of the bow and set-size effects in absolute identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 289–305.

Keppel, G. (1964). Facilitation in short- and long-term retention of paired associates following distributed practice in learning. *Journal of Verbal Learning and Verbal Behavior, 3,* 91–111.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1,* 153–161.

Kieras, D. E., Meyer, D. E., Mueller, S., & Seymour, Y. (1999). Insights into working memory from the perspective of the EPIC architecture for modelling skilled perceptual-motor and cognitive human performance. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 183–223). Cambridge, England: Cambridge University Press.

Kincaid, J. P., & Wickens, D. D. (1970). Temporal gradient of release from proactive inhibition. *Journal of Experimental Psychology, 86,* 313–316.

Koffka, K. (1935). *Principles of Gestalt psychology.* London: Routledge & Kegan Paul.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99,* 22–44.

Kruschke, J. K., & Johansen, M. K. (1999). A model of probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1083–1119.

Lacouture, Y., Grondin, S., & Mori, S. (2001). Absolute identification of temporal intervals: Preliminary data. In E. Sommerfeld, R. Kompass, & T. Lachmann (Eds.), *Proceedings of the Seventeenth Meeting of the International Society of Psychophysics* (pp. 493–498). Berlin, Germany: Pabst Science Publishers.

Lacouture, Y., & Marley, A. A. J. (2004). Choice and response time processes in the identification and categorization of unidimensional stimuli. *Perception & Psychophysics, 66,* 1206–1226.

Lamberts, K. (1995). Categorization under time pressure. *Journal of Experimental Psychology: General, 124,* 161–180.

Laming, D. (1992). Analysis of short-term retention: Models for Brown–Peterson experiments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 1342–1365.

Laming, D. (1999). Testing the idea of distinct storage mechanisms in memory. *International Journal of Psychology, 34,* 419–426.

Laming, D. (2006). Predicting free recalls. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1146–1163.

Lansdale, M. W. (1998). Modeling memory for absolute location. *Psychological Review, 105,* 351–378.

Lee, C. L., & Estes, W. K. (1981). Item and order information in short-term memory: Evidence for multilevel perturbation processes. *Journal of Experimental Psychology: Human Learning and Memory, 7,* 149–169.

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. (2005). Serial recall and presentation schedule: A micro-analysis of local distinctiveness. *Memory, 13,* 283–292.

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

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

Lippman, L. G. (1980). Perceptual isolation and position judgment. *American Journal of Psychology, 93,* 683–695.

Loess, H., & Waugh, N. C. (1967). Short-term memory and inter-trial interval. *Journal of Verbal Learning and Verbal Behavior, 6,* 87–92.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492–527.

Lovett, M. C., Reder, L. M., & Lebiere, C. (1999). Modelling working memory in a unified perspective: An ACT–R perspective. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 135–182). Cambridge, England: Cambridge University Press.

Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 103–190). New York: Wiley.

Maddox, W. T. (1999). On the dangers of averaging across observers when comparing decision bound models and generalized context models of categorization. *Perception & Psychophysics, 61,* 354–374.

Mandler, G., Pearlstone, Z., & Koopmans, H. S. (1969). Effects of organization and semantic similarity on recall and recognition. *Journal of Verbal Learning and Verbal Behavior, 8,* 410–423.

Marcer, D. (1972). Interference and short-term retention. *British Journal of Psychology, 63,* 533–536.

Maylor, E. A. (2002). Serial position effects in semantic memory: Reconstructing the order of verses of hymns. *Psychonomic Bulletin & Review, 9,* 816–820.

Maylor, E. A., Chater, N., & Brown, G. D. A. (2001). Scale invariance in the retrieval of retrospective and prospective memories. *Psychonomic Bulletin & Review, 8,* 162–167.

Maylor, E. A., Vousden, J. I., & Brown, G. D. A. (1999). Adult age differences in short-term memory for serial order: Data and a model. *Psychology and Aging, 14,* 572–594.

McCormack, T., Brown, G. D. A., Maylor, E. A., Richardson, L. B. N., & Darby, R. J. (2002). Effects of aging on absolute identification of duration. *Psychology and Aging, 17,* 363–378.

McCrary, J. W., & Hunter, W. S. (1953). Serial position curves in verbal learning. *Science, 117,* 131–134.

McElree, B. (1996). Accessing short-term memory with semantic and phonological information: A time-course analysis. *Memory & Cognition, 24,* 173–187.

McGeoch, J. A., & Irion, A. L. (1952). *The psychology of human learning.* New York: Longmans.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85,* 207–238.

Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior, 2,* 1–21.

Melton, A. W., & Irwin, J. M. (1940). The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *American Journal of Psychology, 53,* 173–203.

Mensink, G. J., & Raaijmakers, J. G. W. (1988). A model for interference and forgetting. *Psychological Review, 95,* 434–455.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

Miller, G. A., & Nicely, P. A. (1955). An analysis of perceptual confusions among some English consonants. *Journal of the Acoustical Society of America, 27,* 338–352.

Modigliani, V., & Hedges, D. G. (1987). Distributed rehearsals and the primacy effect in single-trial free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 426–436.

Morton, J. (1968). Repeated items and decay in memory. *Psychonomic Science, 10,* 219–220.

Murdock, B. B., Jr. (1960). The distinctiveness of stimuli. *Psychological Review, 67,* 16–31.

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

Murdock, B. B., Jr. (1974). *Human memory: Theory and data.* Potomac, MD: Erlbaum.

Murdock, B., & Metcalfe, J. (1978). Controlled rehearsal in single-trial free recall. *Journal of Verbal Learning and Verbal Behavior, 17,* 309–324.

Muter, P. (1979). Response latencies in discriminations of recency. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 160–169.

Myung, I. J., Kim, C., & Pitt, M. A. (2000). Toward an explanation of the power law artifact: Insights from response surface analysis. *Memory & Cognition, 28,* 832–840.

Myung, I. J., & Pitt, M. A. (1997). Applying Occam's razor in modeling cognition: A Bayesian approach. *Psychonomic Bulletin & Review, 4,* 79–95.

Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory & Cognition, 16,* 343–352.

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

Nairne, J. S. (1991). Positional uncertainty in long-term memory. *Memory & Cognition, 19,* 332–340.

Nairne, J. S. (1992). The loss of positional certainty in long-term memory. *Psychological Science, 3,* 199–202.

Nairne, J. S. (1996). Short-term/working memory. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 101–126). New York: Academic Press.

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., & Kelley, M. R. (1999). Reversing the phonological similarity effect. *Memory & Cognition, 27,* 45–53.

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

Neath, I. (1993a). Contextual and distinctive processes and the serial position functions. *Journal of Memory and Language, 32,* 820–840.

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

Neath, I. (1998). *Human memory: An introduction to research, data, and theory.* Pacific Grove, CA: Brooks/Cole.

Neath, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review, 7,* 403–423.

Neath, I., Bireta, T. J., & Surprenant, A. M. (2003). The time-based word length effect and stimulus set specificity. *Psychonomic Bulletin & Review, 10,* 430–434.

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., & Brown, G. D. A. (2007). Making distinctiveness models of memory distinct. In J. S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger III* (pp. 125–140). New York: Psychology Press.

Neath, I., Brown, G. D. A., McCormack, T., Chater, N., & Freeman, R. (2006). Distinctiveness models of memory and absolute identification:

Evidence for local, not global, effects. *Quarterly Journal of Experimental Psychology, 59,* 121–135.

Neath, I., & Crowder, R. G. (1990). Schedules of presentation and distinctiveness in human memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 316–327.

Neath, I., & Knoedler, A. J. (1994). Distinctiveness and serial position effects in recognition and sentence processing. *Journal of Memory and Language, 33,* 776–795.

Neath, I., & Nairne, J. S. (1995). Word-length effects in immediate memory: Overwriting trace-decay theory. *Psychonomic Bulletin & Review, 2,* 429–441.

Ng, H. L. H. (1996). *Are time dependent oscillators responsible for temporal grouping effects in short-term memory?* Unpublished honors thesis, University of Western Australia, Nedlands, Australia.

Ng, M. L. H., & Maybery, M. T. (2002). Grouping in short-term verbal memory: Is position coded temporally? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 55*(A), 391–424.

Nilsson, L.-G., Wright, E., & Murdock, B. B. (1975). The effects of visual presentation method on single-trial free recall. *Memory & Cognition, 3,* 427–433.

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.

Nosofsky, R. M. (1986). Attention, similarity and the identification–categorization relationship. *Journal of Experimental Psychology: General, 115,* 39–57.

Nosofsky, R. M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology, 43,* 25–53.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104,* 266–300.

Ornstein, P. A., Naus, M. J., & Stone, B. P. (1977). Rehearsal training and developmental differences in memory. *Developmental Psychology, 13,* 15–24.

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

Parducci, A. (1992). Comment on Haubensak's associative theory of judgment. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 310–313.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & Cognition, 17,* 398–422.

Peterson, L. R. (1966). Short-term verbal memory and learning. *Psychological Review, 73,* 193–207.

Peterson, L. R., & Gentile, A. (1965). Proactive interference as a function of time between tests. *Journal of Experimental Psychology, 70,* 473–478.

Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual items. *Journal of Experimental Psychology, 61,* 193–198.

Petrusic, J. M., & Jamieson, D. G. (1978). Differential interpolation effects in free recall. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 101–109.

Pinto, A. C., & Baddeley, A. D. (1991). Where did you park your car? Analysis of a naturalistic long-term recency effect. *European Journal of Cognitive Psychology, 3,* 297–313.

Pollack, I. (1952). The information of elementary auditory displays: I. *Journal of the Acoustical Society of America, 24,* 745–749.

Pollio, H. R., & Deitchman, R. (1964). *The activational characteristics of a serial cognitive structure having oppositional end points.* Mimeographed manuscript, University of Tennessee.

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

Riefer, D. M., & LaMay, M. L. (1998). Memory for common and bizarre stimuli: A storage-retrieval analysis. *Psychonomic Bulletin & Review, 5,* 312–317.

Roediger, H. L., III, & Crowder, R. G. (1976). A serial position effect in recall of United States presidents. *Bulletin of the Psychonomic Society, 8,* 275–278.

Rubin, D. C., Hinton, S., & Wenzel, A. (1999). The precise time course of retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 1161–1176.

Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review, 103,* 734–760.

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

Ryan, J. (1969a). Grouping and short-term memory: Different means and patterns of grouping. *Quarterly Journal of Experimental Psychology, 21,* 137–147.

Ryan, J. (1969b). Temporal grouping, rehearsal, and short-term memory. *Quarterly Journal of Experimental Psychology, 21,* 148–155.

Schneider, W. (1999). Working memory in a multilevel hybrid connectionist control architecture (CAP2). In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 340–374). Cambridge, England: Cambridge University Press.

Schweickert, R., Chen, S., & Poirier, M. (1999). Redintegration and the useful lifetime of the verbal memory representation. *International Journal of Psychology, 34,* 447–453.

Sehulster, J. R. (1989). Content and temporal structure of autobiographical knowledge: Remembering 25 seasons at the Metropolitan Opera. *Memory & Cognition, 17,* 590–606.

Shepard, R. N. (1957). Stimulus and response generalization: A stochastic model relating generalization to distance in psychological space. *Psychometrika, 22,* 325–345.

Shepard, R. N. (1987, September 11). Toward a universal law of generalization for psychological science. *Science, 237,* 1317–1323.

Shiffrin, R. M., & Nosofsky, R. M. (1994). 7 plus or minus 2: A commentary on capacity limitations. *Psychological Review, 101,* 357–361.

Stewart, N., Brown, G. D. A., & Chater, N. (2005). Absolute identification by relative judgment. *Psychological Review, 112,* 881–911.

Surprenant, A. M., Neath, I., & Brown, G. D. A. (2006). Modeling age-related differences in immediate memory using SIMPLE. *Journal of Memory and Language, 55,* 572–586.

Tan, L., & Ward, G. (2000). A recency-based account of the primacy effect in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1589–1625.

Treisman, M. (1985). The magical number seven and some other features of category scaling: Properties of a model for absolute judgment. *Journal of Mathematical Psychology, 29,* 175–230.

Tulving, E. (1966). Subjective organization and effects of repetition in multi-trial free-recall learning. *Journal of Verbal Learning and Verbal Behavior, 5,* 193–197.

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.

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

Underwood, B. J. (1957). Interference and forgetting. *Psychological Review, 64,* 49–60.

Underwood, B. J. (1977). *Temporal codes for memories: Issues and problems.* Hillsdale, NJ: Erlbaum.

Underwood, B. J., & Ekstrand, B. R. (1967). Studies of distributed practice: XXIV. Differentiation and proactive inhibition. *Journal of Experimental Psychology, 74,* 574–580.

Underwood, B. J., & Freund, J. S. (1968). Effect of temporal separation of two tasks on proactive inhibition. *Journal of Experimental Psychology, 78,* 50–54.

Vousden, J., & Brown, G. D. A. (1998). To repeat or not to repeat: The time course of response suppression in sequential behaviour. In J. A. Bullinaria, D. W. Glasspool, & G. Houghton (Eds.), *Proceedings of the Fourth Neural Computation and Psychology Workshop: Connectionist representations* (pp. 301–315). London: Springer-Verlag.

Vousden, J. I., Brown, G. D. A., & Harley, T. A. (2000). Serial control of phonology in speech production: A hierarchical model. *Cognitive Psychology, 41,* 101–175.

Wallace, W. P. (1965). Review of the historical, empirical, and theoretical status of the von Restorff phenomenon. *Psychological Bulletin, 63,* 410–424.

Watkins, M. J., Neath, I., & Sechler, E. S. (1989). Recency effect in recall of a word list when an immediate memory task is performed after each word presentation. *American Journal of Psychology, 102,* 265–270.

Watkins, M. J., & Watkins, O. C. (1974). Processing of recency items for free recall. *Journal of Experimental Psychology, 102,* 488–493.

Wedell, D. H. (1996). A constructive–associative model of the contextual dependence of unidimensional similarity. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 634–661.

Wickelgren, W. A. (1967). Rehearsal grouping and hierarchical organisation of serial position cues in short-term memory. *Quarterly Journal of Experimental Psychology, 19,* 97–102.

Wickelgren, W. A. (1973). The long and the short of memory. *Psychological Bulletin, 80,* 425–438.

Wickelgren, W. A., Corbett, A. T., & Dosher, B. A. (1980). Priming and retrieval from short-term memory: A speed accuracy trade-off analysis. *Journal of Verbal Learning and Verbal Behavior, 19,* 387–404.

Wickens, D. D., Born, D. G., & Allen, C. K. (1963). Proactive inhibition and item similarity in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 2,* 440–445.

Wickens, T. D. (1999). Measuring the time course of retention. In C. Izawa (Ed.), *On human memory: Evolution, progress, and reflections on the 30th anniversary of the Atkinson–Shiffrin model* (pp. 245–266). Mahwah, NJ: Erlbaum.

Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting functions. *Psychological Science, 2,* 409–415.

Wixted, J. T., & Ebbesen, E. B. (1997). Genuine power curves in forgetting: A quantitative analysis of individual subject forgetting functions. *Memory & Cognition, 25,* 731–739.

Wixted, J. T., & McDowell, J. J. (1989). Contributions to the functional analysis of single-trial free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 685–697.

Wright, A. A., Cook, R. G., Rivera, J. J., Shyan, M. R., Neiworth, J. J., & Jitsumori, M. (1990). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 1043–1059.

Appendix

llustrative Implementation of SIMPLE

In the first part of this appendix, we note that the similarity between items' memory locations can be expressed either in terms of temporal distance ratios or as a negative exponential function of the separation of the locations along a logarithmically transformed temporal distance dimension; the equivalence is straightforward. In the second part, we give a simple worked example of the calculations underpinning the model for free recall and serial recall.

The similarity between two memory locations is

$$\eta_{i,j} = e^{-c|M_i - M_j|},$$

and using $M_x = \log(T_x)$, where T_x is the temporal distance of item x, this becomes

$$\eta_{i,j} = e^{-c|\log(T_i) - \log(T_j)|} = e^{-c|\log(T_i/T_j)|}.$$

$|\log(T_i/T_j)| = \log(T_i/T_j)$ if $T_i \geq T_j$, and $|\log(T_i/T_j)| = \log(T_j/T_i)$ if $T_i < T_j$, so

$$\eta_{i,j} = (T_j/T_i)^c \qquad \text{if } T_i \geq T_j$$
$$\eta_{i,j} = (T_i/T_j)^c \qquad \text{if } T_i < T_j.$$

The interpretation of this is straightforward. Given temporal distances of T_i and T_j, say 5 and 15, the similarity between them is the smaller value divided by the larger (here $5/15 = 1/3$) raised to the power c. If the values are separated by a large ratio (e.g., the temporal distances are 1 and 10), the similarity will be small ($.1^c$); conversely, if they are identical, their similarity will be maximal, at 1, equal to their ratio. Let us define a function Ratio(x, y), which divides the smaller of x and y by the larger. Then we can write

$$\eta_{i,j} = \text{Ratio}(T_i, T_j)^c.$$

In other words, the similarity of two memory values is some power of the ratio of their temporal distances.

We now illustrate the case of free and serial recall with a worked numerical example involving four items presented at a rate of one per second, with recall commencing 1 s after the offset of the final list item.

The first step is to construct the temporal distances (ages) of the item offsets at the point that they are recalled. The time of recall for successive items will depend on the task. For free recall, because recall dynamics are often not known, we will assume that every item is recalled 4 s after the offset of the final list item. For serial recall, we will assume that each successive item recall takes 2 s to recall.

Calculations are illustrated in Table A1. Rows represent successive recall attempts, so the untransformed temporal distances of all list items at the time of the attempt to recall each item will be as shown in the table (top two sets of 16 numbers). In the case of serial recall, the nature of the task requires that the first item be recalled on the first recall attempt, the second item on the second recall attempt, and so on. In the case of free recall, there is no such requirement.

We now consider each case separately. For free recall, each item

(*Appendix continues*)

Table A1
Illustration of Calculations of the SIMPLE Model

Recall attempt/ cued item location	Free recall 1	Free recall 2	Free recall 3	Free recall 4	Serial recall 1	Serial recall 2	Serial recall 3	Serial recall 4
				Temporal distances of items at time of retrieval				
1	7	6	5	4	4	3	2	1
2	7	6	5	4	6	5	4	3
3	7	6	5	4	8	7	6	5
4	7	6	5	4	10	9	8	7
				Discriminabilities of items				
1	.69	.23	.07	.01	.88	.12	.01	.00
2	.20	.60	.17	.03	.18	.66	.14	.02
3	.06	.18	.63	.13	.08	.19	.57	.16
4	.02	.05	.16	.78	.05	.10	.24	.61
				Item retrieval probabilities				
1	.67	.05	.01	.01	.90	.02	.01	.00
2	.04	.49	.03	.01	.03	.62	.02	.01
3	.01	.03	.56	.02	.01	.04	.44	.03
4	.01	.01	.03	.80	.01	.02	.05	.51
				Serial position curves				
	.73	.59	.64	.85	.90	.62	.44	.51

i will have a certain discriminability in memory at the time of each retrieval attempt j. In other words, given the memory location of item j (i.e., T_j), each item will have a certain discriminability. These are given by

$$D_i|T_j = \frac{\eta_{i,j}}{\sum_{k=1}^{n}(\eta_{j,k})},$$

where

$$\eta_{i,j} = e^{-c|\log(T_i)-\log(T_j)|}.$$

With c set arbitrarily to 7.0 for the purposes of illustration, the resulting discriminabilities of each item, at each recall-attempt/cued-item location, will be as shown in Table A1 (left set of 16 numbers; middle of table). Note that row totals must sum to 1, but column totals need not.

The next step is to transform item discriminabilities into recall probabilities, using the thresholding function:

$$P(R_i|D_i) = \frac{1}{1 + e^{-s(D_i-t)}},$$

where R_i is the probability of recalling item i on a given retrieval attempt, and s and t are the slope and threshold parameters, respectively. Using arbitrary parameter values $t = 0.6$ and $s = 8$, application of the thresholding function produces the item retrieval probabilities shown in the table (lowest block of 16). For example, given the temporal location of the third item, there is a .56 probability of recalling Item 3 and a .03 probability of recalling Item 2. In free recall an item can be counted as correct even if it is recalled as a result of the attempt to recall some other item. Therefore, the free recall serial position curve is obtained by summing the columns of retrieval probabilities, and the resulting serial position curve is shown as the last row of the table. If the recall probability for a given item is greater than 1.0, it is set to 1.0. Note the large recency evident in the serial position curve, along with a small amount of primacy.

Next, we consider serial recall. Note the changing temporal perspective of recall; when the first item must be recalled it is 4 s in the past, but when the fourth item must be recalled it is 7 s in the past and hence less temporally distinctive. Item discriminabilities and item retrieval probabilities are calculated in just the same way as for the free recall case, using the same parameter values for illustration. The thresholding mechanism could cause some probabilities to become greater than 1.0; in such cases they are set to 1.0.

In the case of serial recall, an item will be scored as correct only if it is recalled at the appropriate attempt. The serial position curve, as shown in the bottom row of the table, is therefore given by the diagonal of the matrix of retrieval probabilities (i.e., Item 1 is counted as correct only if it is recalled during the attempt to recall the first item, and so on). The slow time-course of recall relative to list presentation in this time-only version of the model produces extended primacy in the serial position curve for forward serial recall; see the main text for discussion of output interference as an alternative source of such primacy.

Received September 11, 2006
Revision received March 14, 2007
Accepted March 15, 2007 ■