# Why do participants initiate free recall of short lists of words with the first list item? Toward a general episodic memory explanation.

Jessica Spurgeon, Geoff Ward, & William J. Matthews

University of Essex

Correspondence to:

Jessica Spurgeon or Geoff Ward,

Department of Psychology,

University of Essex,

Wivenhoe Park,

Colchester, Essex

CO4 3SQ, UK

Telephone: +44 1206 873799 or +44 1206 873800

Fax: +44 1206 873590

E-mail: jsmithy@essex.ac.uk or gdward@essex.ac.uk

Running header: the free recall of short word lists

# Abstract

Participants who are presented with a short list of words for immediate free recall (IFR) show a strong tendency to initiate their recall with the first list item and then proceed in forward serial order. We report two experiments that examined whether this tendency was underpinned by a short-term memory store, of the type that is argued by some to underpin recency effects in IFR. In Experiment 1, we presented three groups of participants with lists of between 2 and 12 words for IFR, delayed free recall (DFR), and continuous-distractor free recall (CDFR). The to-be-remembered words were simultaneously spoken and presented visually, and the distractor task involved silently solving a series of self-paced, visually-presented mathematical equations (e.g., “3+2+4=?”). The tendency to initiate recall at the start of short lists was greatest in IFR, but was also present in the two other recall conditions. This finding was replicated in Experiment 2, where the to-be-remembered items were presented visually in silence and the participants spoke aloud their answers to computer-paced mathematical equations. Our results necessitate that a short-term buffer cannot be fully responsible for the tendency to initiate recall from the beginning of a short list, but rather suggest that the tendency represents a general property of episodic memory that occurs across a range of timescales.


*Keywords*: free recall; serial recall; output order; short-term memory; episodic memory

215 words

Ward, Tan, and Grenfell-Essam (2010) have recently reported a novel experimental finding concerning the immediate free recall (IFR) of short lists: when participants were asked to recall a short list of words *in any order* they often responded i*n forward serial order.* That is, if asked to recall in any order “dog, house, man, stairs”, they often recalled “dog, house, man, stairs”, even though there was no formal requirement to perform immediate serial recall (ISR). Ward *et al.* argued that this phenomenon of initiating IFR of short lists with the first list item and proceeding in forward serial order demonstrated that the IFR and ISR tasks were in fact more similar than had often been previously assumed. Indeed, a number of similarities were observed across the two tasks, when IFR and ISR were compared at the same list lengths using the same scoring systems (for related data, see Bhatarah, Ward, Smith, & Hayes, 2009; Grenfell-Essam & Ward, 2012).

The aim of this current paper is not so much to encourage further the theoretical integration of the ISR and IFR literatures (but for informative recent reviews, see Farrell, 2012; Grenfell-Essam & Ward, 2012; Hurlstone, Hitch, & Baddeley, 2014; Kahana, 2012b), but to consider the more fundamental question of why participants initiate IFR of a short list of words with the first list item when there is no requirement to do so. As identified by Ward *et al.* (2010), the tendency to initiate recall with the first list item represents a problem for unitary accounts of IFR that would otherwise emphasise the heightened accessibility of the most recent items in the list. In the remainder of this introduction, we briefly review the current debate between dual-store and unitary accounts of IFR, we discuss why unitary accounts of IFR may have particular difficulty in accommodating the Ward *et al.* findings, and we consider the possibility that this phenomenon of interest may be explained within a dual-store account of IFR. Finally, we outline a pair of studies designed to test a short-term memory store (STS) explanation of the phenomenon. To anticipate our findings, we show that the tendency to initiate IFR of short lists with the first list item occurs (albeit at a reduced rate) even in methodologies that are designed to render a short-term memory buffer store inoperative.

## The current debate between dual-store and unitary accounts of IFR

Early accounts of IFR were primarily concerned with explaining the characteristic U-shaped serial position curves associated with the task. When participants are presented with lists of between 10 and 40 words for IFR, they typically show large and extended recall advantages for words presented towards the end of the list, as well as more modest recall advantages for words presented towards the start of the list (the recency and primacy effects, respectively, Murdock, 1962).

Early dual-store accounts (e.g., Atkinson & Shiffrin, 1968, 1971; Glanzer, 1972; Raaijmakers & Shiffrin, 1981; Waugh & Norman, 1965) proposed that the recency effects in IFR reflected the direct output from a highly accessible, yet limited-capacity, STS or buffer. It was argued that the words towards the end of the list were those most likely to be in the STS at test, and these words were output first, in order to prevent them from being displaced by the recall of other list items. After the STS buffer had been emptied, words would then be retrieved from the long-term memory store (LTS), and recall from LTS would be affected by factors such as the length of time spent in STS.

Dual-store explanations were well supported by the findings from a variant of the free recall task called delayed free recall (DFR) in which a period of rehearsal-preventing distractor activity was inserted between the end of the list and recall. Typical distractor tasks used in DFR have included backwards counting (Gardiner, Thompson & Maskarinec, 1974; Martin & Jones, 1979; Postman & Phillips, 1965), counting aloud (Glanzer & Cunitz, 1966; Raymond, 1969), solving simple mathematical equations (Greene, 1986; Howard & Kahana, 1999), digit shadowing (Gardiner et al., 1974), and vowel detection (Martin & Jones, 1979). Under these conditions, it was assumed that the filled distractor activity would displace the recency items from STS. Consistent with such an interpretation, whilst the primacy effect in DFR is typically unaffected relative to IFR, the recency effect is greatly reduced, if not eliminated, by a filled delay (e.g., Howard & Kahana, 1999; Glanzer & Cunitz, 1966; Glenberg, Bradley, Kraus & Ranzaglia, 1983; Greene, 1986; Lehman & Malmberg, 2013; Postman & Phillips, 1965; Raymond, 1969). Dual-store accounts of IFR were further supported by the complementary finding that a number of variables associated with long-term learning, such as word frequency (e.g., Sumby, 1963), list length (e.g., Murdock, 1962), and presentation rate (e.g., Glanzer & Cunitz, 1966) were shown to affect the early and middle list positions but not the recency positions, a result that was interpreted as demonstrating that these variables selectively affected LTS but not STS (Glanzer, 1972).

However, one potential difficulty for early dual-store accounts of IFR is that long-term recency effects were observed under conditions in which an STS explanation was untenable. Recency effects were observed in the recall of real-world events that occurred over a timescale of days, weeks, and months (e.g., Baddeley & Hitch, 1974; Pinto & Baddeley, 1991) and were also observed in the laboratory, using the continual distractor free recall (CDFR) task, in which a filled period of distractor activity was inserted after each and every list item including the last (e.g., Bjork & Whitten, 1974; Glenberg et al., 1980; Greene, 1986; Nairne, Neath, Serra & Byun, 1997). This finding has been extensively studied and the recency effect in CDFR widely replicated (e.g., Bhatarah, Ward & Tan, 2006; Bjork & Whitten, 1974; Davelaar et al., 2005; Gardiner & Gregg, 1979; Glenberg et al., 1983; Glenberg & Swanson, 1986; Howard & Kahana, 1999; Koppenaal & Glanzer, 1990; Lehman & Malmberg, 2013; Neath, 1993; Poltrock & MacLeod, 1977; Tzeng, 1973; Watkins, Neath & Sechler, 1989; Whitten, 1978).

A second potential difficulty for early dual-store accounts of IFR concerns the output order of the recalled items in tests of free recall. Participants recalling a list of 10 to 40 words typically initiate recall with one of the last few words (e.g., Beaman & Morton, 2000; Farrell, 2010; Hogan, 1975; Howard & Kahana, 1999; Laming, 1999) and there is a tendency for successive recalls to be words that were presented in neighbouring list positions (the temporal contiguity effect, Kahana, 1996). There is a particularly strong tendency for output to proceed in forward serial order. In principle, temporal contiguity effects in IFR could be readily explained by dual-store accounts, because co-occurrence in the STS can lead to increased association between neighbouring items in LTS. The difficulty for dual store accounts is that temporal contiguity effects can also be demonstrated under CDFR conditions (Howard & Kahana, 1999) and over far longer time-scales (e.g. Howard, Youker, & Venkatadass, 2008; Moreton & Ward, 2010). These long-term contiguity effects have been assumed by some unitary theorists to demonstrate that temporal contiguity effects are timescale-invariant (or timescale-similar), and at the very least not always attributable to STS.

One reaction to these long-term recency and long-term contiguity effects has been to abandon the distinction between STS and LTS. Unitary accounts of IFR assume that the same memory mechanisms underpin recall from all serial positions. These accounts typically assume that the to-be-remembered list items are represented along a continuum of episodic memory, with list items positioned along a temporal dimension (e.g., Brown, Neath & Chater, 2007; Glenberg, 1987; Glenberg & Swanson, 1986) or associated with a drifting temporal context (e.g., Howard & Kahana, 2002; Polyn, Norman & Kahana, 2009; Sederberg, Howard, & Kahana, 2008; Tan & Ward, 2000). Unitary accounts assume that the most recent items tend to be the most accessible list items; owing to their greater temporal distinctiveness (e.g., Brown et al., 2007; Tan & Ward, 2000) or because the most recent items are associated with temporal contexts that most closely resemble the temporal context at test (e.g., Howard & Kahana, 1999). Temporal context accounts can predict temporal contiguity effects at all timescales (e.g., Howard & Kahana, 2002) because the partially retrieved temporal context associated with each retrieved item can act as a retrieval cue to help access the similar temporal contexts of neighbouring items.

An alternative reaction has been to acknowledge the necessity for some long-term mechanism to underpin recency and contiguity effects in CDFR, but to further assume that some additional short-term memory mechanism may also be necessary to explain the enhanced recency effects and enhanced temporal contiguity effects that are typically observed with the first few outputs in IFR (e.g. Davelaar et al., 2005; Farrell, 2010; Lehman & Malmberg, 2013; Unsworth & Engle, 2007).

Resolving this debate is potentially difficult because both short-term and long-term recency and contiguity mechanisms are assumed to produce qualitatively similar patterns of effects. It is perhaps unsurprising, therefore, that there remains continued controversy over whether short-term memory is necessary to explain immediate memory phenomena. Recent commentaries in favour of short-term memory include: Davelaar, Usher, Haarmann, & Goshen-Gottstein (2008), Thorn and Page (2009), and Usher, Davelaar, Haarmann, and Goshen-Gottstein (2008), whereas unitary explanations can be found in Neath & Brown (2006), Sederberg et al. (2008), and Surprenant and Neath (2009).

## The difficulty that the Phenomenon of Interest poses to unitary accounts

One striking feature of the Ward *et al.* (2010) finding is that the tendency to initiate IFR of short lists with the first list item appears to be qualitatively different from that predicted by the recency-dominated, unitary accounts of IFR (e.g., Brown et al., 2007; Glenberg & Swanson, 1986; Greene, 1986; Howard & Kahana, 2002; Tan & Ward, 2000; Ward, 2002). Moreover, following the initial recall of the first list item, recall tends to proceed in a forward serial order, such that participants perform IFR of short lists of words in an “ISR-like” manner (Ward et al., 2010), as evidenced by *primacy-dominated* serial position curves when recall is scored using serial recall (SR) scoring. As already noted, unitary accounts correctly predict that participants will initiate IFR of longer lists of words with one of the last few words in the list (most commonly the very last word), but it is difficult to understand why this prediction should be so dramatically different at shorter list lengths. The Ward *et al.* (2010) finding is therefore difficult to reconcile with these unitary accounts.

Unitary accounts have often relied upon rehearsal to explain primacy effects in IFR (e.g., Laming, 2006, 2008, 2009, 2010; Tan & Ward, 2000; Ward, 2002; Ward & Tan, 2004; Ward, Woodward, Stevens & Stinson, 2003). Consistent with these explanations, at slow rates, early list items are rehearsed more often (e.g., Rundus, 1971; Tan & Ward, 2000), they are distributed more widely throughout the list (Modigliani & Hedges, 1987; Tan & Ward, 2000), and they are rehearsed to more recent list positions (e.g., Brodie & Murdock, 1977; Tan & Ward, 2000). Moreover, the rehearsal order is a good predictor of later recall order (Laming, 2006, 2008, 2009, 2010; Ward, et al., 2003). It is clear that rehearsal can, in principle, allow recency-based accounts to explain why early list items are far better recalled than would otherwise be predicted from their nominal serial position.

However, Grenfell-Essam, Ward, and Tan (2013) have recently shown that the tendency to initiate IFR of short lists of words with the first list item cannot be attributed to selective rehearsal. Grenfell-Essam et al. observed that the tendency was unaffected by a doubling of the presentation rate from 1 to 2 words per second, and was still present (albeit reduced) when the spoken stimuli were presented under concurrent articulation (CA). A rehearsal interpretation has been further undermined by the recent findings of Spurgeon, Ward, and Matthews (2014) who have shown that the finding is even observed for visually-presented words presented under CA, conditions where it is usually assumed that the visual stimuli cannot even be phonologically recoded. It would seem therefore that there is something special about the first list item in short lists that is as yet not readily identified in many recency-based unitary accounts of IFR.

Alternatively, unitary accounts of memory that are based on temporal distinctiveness, (e.g., Brown *et al.*, 2007) could attribute the heightened accessibility of the first list item to the increased distinctiveness of the first list item (since the first list item has no preceding list items and so is in a less temporally-crowded region of space). Although temporal distinctiveness accounts do predict "edge" effects for the first item, these accounts cannot explain why the first list item of a short list is so much more accessible than the most distinctive, final list item in an immediate test.

## A dual-store explanation of the phenomenon of interest

A contrary proposition is that the tendency to initiate IFR of short lists of words with the first list item could reflect the existence of a short-term buffer store, of the type that is argued by some dual-store theorists to underpin recency effects in IFR. Output from this STS would normally be evidenced by strong recency effects at conventional list lengths, but if the list was sufficiently short (e.g., lists of only 3 or 4 items), then one might imagine that this same STS buffer memory might be used to output the entire list in order, starting with the very first list item.

As a concrete example, imagine a most basic account of STS that is assumed to consist of say, three or four "slots" arranged for the sake of exposition, from left to right. Let us imagine that the slots when empty are filled up in order from left to right, and that retrieval from STS is again always from left to right. If one assumes that when the STS is full, each additional new item displaces an existing item in STS at random, then one has a very simplistic STS account of first response, that correctly predicts (1) the forward serial recall of very short lists, (2) the negligible tendency to ever initiate recall with the second or third list item, (3) the decreasing tendency to initiate recall with the first list item with increasing list length, (4) the increasing tendency to initiate recall with one of the last four list items with increasing list length, and (5) a recency gradient in the probability of first response at long lists, with the most likely list item to be output first being the very last list item. Clearly, this simplistic account of STS has a number of important deficiencies, but it is useful to illustrate that an STS that contributes to recency effects at long lists, could at least in principle be responsible for the tendency to initiate IFR of short lists with the first list item.

## Experiment 1

Experiment 1 examines whether a STS that is sometimes assumed to account for recency effects in IFR (Atkinson & Shiffrin, 1968; Davelaar et al., 2005; Farrell, 2010; Lehman & Malmberg, 2013; Raaijmakers & Shiffrin, 1981; Unsworth & Engle, 2007) could also account for the tendency to initiate IFR of a short list of words with the first item (Grenfell-Essam & Ward, 2012, Grenfell-Essam et al., 2013; Spurgeon et al., 2014; Ward et al., 2010). To this end, Experiment 1 examines the output order observed in three variants of free recall that were each performed with lists of between two and twelve words. The three variants of free recall were IFR, DFR, and CDFR. The assumption was that a period of filled distractor activity that occurred at the end of each list (DFR) or that occurred after each and every list item including the last (CDFR) should overwrite the contents of STS at test (DFR), and at test and during encoding (CDFR). The main prediction was that the tendency to initiate IFR of short lists of words with the first list item should be all but eliminated in the DFR and CDFR conditions, if the tendency was underpinned solely by the retrieval of the contents of STS at test.

Furthermore, as noted by Glenberg and Swanson (1986), we assumed that the presence of distractors during the inter-presentation intervals in CDFR would discourage the formation of associations between successive list items and thus prevent their rehearsal into a sequence, which remains possible in both IFR and DFR. Thus, if the tendency to initiate IFR of short lists with the first list item occurs in DFR but not CDFR, then this would suggest that the tendency arises, at least in part, from the opportunity to form associations between the list items in STS at encoding, during the presentation of the list.

In Experiment 1, the to-be-remembered words in all three conditions were presented visually and simultaneously spoken to the participant, whereas the distractor task used in DFR and CDFR consisted of solving a series of self-paced visual mathematics equations.

## Method

**Participants.** Sixty psychology students from the University of Essex participated in exchange for course credit.

**Materials and apparatus.** The materials consisted of a pool of 478 words drawn from the Toronto Noun Pool (Friendly, Franklin, Hoffman, & Rubin, 1982). Subsets of 158 words were randomly selected for each participant to be the to-be-remembered list items for the experiment. Furthermore, a set of 84 simple mathematical equations was constructed, which consisted of all possible additions of three single-digit numbers (e.g., “2 + 1 + 5 = ?”), for which the result was always a positive value between 3 and 9. Using the application, Supercard, all words and equations were presented visually in the centre of an Apple Macintosh computer screen. At the same time as each word was presented visually, each word was also spoken using the digitised voice files of the Toronto Noun Pool (obtained from Michael Kahana's Computational Memory Laboratory website 2012, http://memory.psych.upenn.edu/WordPools). Each participant was provided with a response booklet consisting of 26 response grids, each of which contained two columns and 12 rows. The first column of each grid was narrow and contained the numbers 1-12 in ascending order. The second column was wider to allow room for participants to write down their responses.

**Design.** The experiment used a mixed design. The between-subjects variable was the variant of the task with three levels (IFR, DFR, or CDFR). There were two within-subjects variables: list length with six levels (2, 3, 4, 6, 9, or 12), and serial position (SP) with up to twelve levels (SP 1-12). The dependent variables were the proportion of words recalled in any order irrespective of output position (free recall, FR scoring) and also the proportion of words recalled in the same output (response grid) position as input serial position (SR scoring). SR scoring was examined because it provides the most conventional measure of forward serial order recall.

**Procedure.** Participants were tested individually and were informed that they would be shown 2 practice lists, followed by 24 experimental lists of words where they must try to remember as many words as possible, in any order that they wished. Participants performing DFR and CDFR were told they would also be required to perform an arithmetic task which consisted of solving a number of mathematical equations either after the presentation of the final word (DFR) or after the presentation of each and every word (CDFR). These participants were told that both the memory and arithmetic tasks were important, and that their answers to the arithmetic task would be recorded and subsequently checked. The practice lists were of 7 words. The 24 experimental trials were divided into two blocks of 12 trials. In each block, participants received two trials at each of the 6 list lengths, and the order of these trials within each block was randomised so that participants were not aware of the length of the lists in advance of its presentation. The words were randomly allocated on each trial and no items were repeated across lists, such that no word appeared more than once to each participant.

Each trial started with a warning tone and a fixation cross, followed after 1 second by a sequence of between 2 and 12 words that were simultaneously spoken by the computer and presented visually in the centre of the computer screen, during which participants remained silent. The words were displayed for 0.75 seconds each with an additional 0.25 seconds inter-stimulus interval during which the computer screen was blank. For those participants performing IFR and DFR, the next word was presented immediately after the preceding word. After the presentation of the final word, those participants performing IFR began recall immediately, whereas those participants performing DFR had a filled retention interval of 15 seconds where they performed the arithmetic task before they began recall. For those participants performing CDFR, there was a filled inter-presentation interval (IPI) of 15 seconds following each and every word where they performed the arithmetic task prior to the presentation of the subsequent word, or in the case of the last word, prior to beginning recall.

The arithmetic task performed by those in the DFR and CDFR groups consisted of solving simple mathematical equations. The equations were randomly sampled from the pool of 84 possible equations of the type described above and were presented silently in the centre of the computer screen. Participants were required to read the equations silently and respond with their answers as quickly and as accurately as possible by pressing the appropriate key (3-9) on the computer keyboard. Each equation was displayed individually and remained on the screen either until an answer was provided or until the distractor task time was up. Once an answer had been provided, there was an interval of 0.25 seconds prior to the onset of the next equation. Once the 15 seconds was up, participants were presented with the next word in the list, or the recall period began.

At the start of each recall period, an empty grid appeared on the screen that contained the same number of numbered rows as there had been words on the list, thereby indicating at a glance the list length of the current trial. Participants wrote down as many words as they could remember on their response sheets (which always contained 12 rows) in any temporal order that they wished, and filled their response grids from the top of the grid. There was no time limit placed on the recall period: participants finished recall once they felt like they had remembered all the words they could, and then they could begin the next trial.

## Results

**Performance on distractor tasks.** The average percentage of equations solved correctly during the mathematics task was 93% in the CDFR condition (with a range of 81-97%) and 94% in the DFR condition (with a range of 86-100%).

**Proportion of words recalled**. The left-hand panels of Figure 1 shows the proportion of words recalled for the three tasks using FR scoring (Figure 1A) and SR scoring (Figure

1C). Table 1 summarises the findings of a pair of 3 (task: IFR, DFR and CDFR) x 6 (list length: 2, 3, 4, 6, 9 and 12) mixed ANOVAs on the proportion of words recalled for both FR and SR scoring.

----------------------------------------

--Figure 1 about here--

----------------------------------------

----------------------------------------

--Table 1 about here--

----------------------------------------

Considering first the data using FR scoring, there was a significant main effect of list length showing better performance at shorter list lengths. Specifically, as the list length increased, so the proportion of words recalled decreased monotonically. There was a significant main effect of task showing that significantly fewer words were recalled in DFR, but the proportion of words recalled in IFR and CDFR did not differ significantly. The significant two-way interaction between task and list length revealed superior IFR and CDFR performance relative to DFR for list length 6; IFR was also better than CDFR at list lengths 2, 3 and 4; CDFR was better than both IFR and DFR at list length 12; and there were no task differences at list length 9.

Considering next the same data using SR scoring, there was a significant main effect of list length showing better performance at shorter list lengths. Specifically, as the list length increased, so the proportion of words recalled in serial order decreased monotonically. There was a significant main effect of task, demonstrating that a greater proportion of words were recalled in serial order in IFR, but the degree of forward order recall did not differ significantly between CDFR and DFR. The two-way interaction between task and list length revealed that IFR was superior to both CDFR and DFR for list lengths 2, 3 and 4; IFR and CDFR did not differ significantly but both were better than DFR at list lengths 9 and 12, and there were no significant task differences at list length 6.

**Analyses of serial position curves of all the data.** Figure 2 shows the serial position curves for each of the three tasks. The left-hand panels show data using FR scoring, with IFR, DFR, and CDFR shown in Figures 2A, 2C, and 2E, respectively. The right-hand panels show the same data using SR scoring, with IFR, DFR, and CDFR shown in Figures 2B, 2D, and 2F, respectively.

----------------------------------------

--Figure 2 about here--

----------------------------------------

Full statistical analyses of the effects of serial position at each list length and task separately on the proportion of correctly recalled words can be found in Appendix A1 (FR scoring), A2 (SR scoring) and A3 (pairwise comparisons). To summarise the general trends, there were similarities when data were compared across tasks, list lengths and scoring methods. Considering first the data with FR scoring, the curves changed in similar ways with increasing list length. Performance was close to ceiling for the very short list lengths, but as the list length was increased, so there were primacy and recency effects at shorter list lengths, and then reduced primacy and increased recency at longer list lengths. There were more words recalled in IFR relative to CDFR, and more words recalled in CDFR relative to DFR, and this was mainly due to the recency effect being strongest in IFR and weakest in DFR.

Considering next the data with SR scoring, there was a similar effect of increasing list length for all tasks. Again, there were more words recalled in IFR relative to CDFR, and more words recalled in CDFR relative to DFR. However, there were no recency effects using SR scoring, a finding that can be readily understood if one considers that participants often recalled the recency items, but they rarely, if ever, wrote these words down in the final output positions of the response grid, resulting in extended primacy.

**The probability of first response (P[FR]) data.** The left-hand panels of Figure 3 show the proportion of trials in which words from different list positions were recalled first for each list length for IFR, DFR, and CDFR shown in Figure 3A, 3C, and 3E, respectively[1]. Following Ward *et al.* (2010), each panel of Figure 4 collapses the initial outputs of a given condition into one of four categories: “SP1” - those trials that started with the first word in the list (that is, the word presented in SP 1), “Last 4”- those trials that started with one of the last four list items[2], “Other” - those trials that started with any of the other list items, and “Void / Error” –those trials in which either nothing was recalled on that trial or where recall began with an intrusion. As can be observed, for all three tasks, at short list lengths, recall was most likely to be initiated with the first word in the list. As list length increased, this tendency decreased and there was a complimentary increase in the tendency to initiate recall with one of the last four words. This resulted in a change in the modal response from ‘SP1’ to ‘Last 4’ at list length 7 or 8 for IFR, reducing to list length 4 for DFR, and reducing still further to list length 3 or 4 for CDFR.

----------------------------------------

--Figure 3 about here--

---------------------------------------

Table 1 summarises the findings of a pair of 3 (task: IFR, DFR, and CDFR) x 6 (list length: 2, 3, 4, 6, 9, and 12) mixed ANOVAs conducted on the proportion of trials that were initiated with the word from SP 1 and the proportion of trials that were initiated with one of the last four words.

We will consider first the proportion of trials in which participants initiated recall with the first list item (that is, when P[FR=SP1]). The significant main effects of task and list length revealed a greater tendency at shorter list lengths and a greater tendency in IFR compared to the other two tasks (DFR and CDFR did not differ significantly). The two-way interaction between task and list length revealed a superior tendency for initiating recall with the first list item in IFR at list lengths 2, 3 and 4. At list lengths 6 and 12, CDFR did not differ significantly from IFR but both were superior to DFR. Finally, at list length 9, there were no significant differences between the tasks.

We will now consider the proportion of trials in which participants initiated recall with one of the last four list items (that is, when P[FR=Last 4]). The significant main effects of task and list length revealed an increased tendency to initiate recall with one of the last four words with increasing list length and also under CDFR and DFR relative to IFR (CDFR and DFR did not differ significantly). The two-way interaction between task and list length revealed this reduction in IFR at list lengths 2, 3 and 4; but there were no significant differences between the tasks at list lengths 6, 9, and 12.

## Discussion

In Experiment 1, we examined whether the tendency to initiate IFR of a short list of words was underpinned by the contents of a limited-capacity STS. This was determined by comparing the output order observed in IFR (where list items could be rehearsed and maintained in STS) with the output order observed in DFR (where list items could be rehearsed in STS, but where the contents of STS would be overwritten by the intervening distractor activity at test) and the output order observed in CDFR (where the contents of STS would be occupied by distractor activity during the interval following each and every list item, including the last). We found that the tendency to initiate recall with the first list item was greatest in IFR, but the tendency was also present (albeit reduced) in both DFR and CDFR, conditions in which an explanation based on the direct output from STS was untenable.

Our results showed some commonalities across the tasks. The proportion of words recalled decreased with increasing list length in all three tasks, and the patterns of recall differed in similar ways in all three tasks as the list length increased. All three tasks showed an increased tendency for serial ordered recall with shorter list lengths. By contrast, all three tasks showed an increased tendency for recall of end of list items with longer list lengths.

These findings confirm prior research that showed extended recency effects present in IFR and CDFR (Bhatarah et al., 2006; Bjork & Whitten, 1974; Davelaar et al., 2005; Gardiner & Gregg, 1979; Glenberg et al., 1983; Glenberg & Swanson, 1986; Howard & Kahana, 1999; Koppenaal & Glanzer, 1990; Lehman & Malmberg, 2013; Neath, 1993; Poltrock & MacLeod, 1977; Tzeng, 1973; Watkins et al., 1989; Whitten, 1978) and are consistent with a reduction in recency effects in DFR (Howard & Kahana, 1999; Glanzer & Cunitz, 1966; Glenberg et al., 1983; Greene, 1986; Lehman & Malmberg, 2013; Postman & Phillips, 1965; Raymond, 1969). These findings are also consistent with prior research demonstrating temporal contiguity effects in IFR and CDFR (e.g. Bhatarah et al., 2006; Howard & Kahana, 1999).

However, one potential concern regarding our data is that although we obtained a reduction in recency in DFR relative to IFR and CDFR, the magnitude of our recency effects in DFR is far greater than in many prior DFR experiments, some of which show that recency has been effectively eliminated when a studied list is followed by a filled period of distractor activity. A corresponding concern, therefore, was that our distractor conditions had not been totally effective at displacing list items from the contents of STS, such that the forward-ordered tendency to recall in short lists in IFR, DFR, and CDFR could still be attributable to the direct output from STS. Experiment 2 sought to allay this potential concern.

## Experiment 2

We were surprised by the degree of recency in DFR in Experiment 1, and compared our methodology carefully with many other DFR and CDFR experiments. In line with other experimenters (Davelaar et al., 2005; Greene, 1986; Howard & Kahana, 1999), we had used mathematical puzzles of the sort "1 + 4 +3 = ?" as our distractor activity, but whereas these prior studies had eliminated recency in DFR, we had only attenuated recency.

We identified two potential differences between our experiment and many other earlier studies. First, we had allowed participants to perform the 15 seconds worth of mathematical equations in Experiment 1 at a self-paced rate. Our participants had solved an average of 6 mathematical problems during this time, taking the equivalent of 2.5 seconds per equation. By contrast, in a preliminary experiment conducted by Greene (1986), who specifically checked for the absence of recency in DFR, the participants were put under time pressure to solve the maths equations, and the rate was experimenter-controlled at 2 seconds per equation. A similar time pressure has been adopted in a many other CDFR studies (Bjork & Whitten, 1974; Glenberg et al., 1980, 1983; Glenberg & Swanson, 1986; Koppenaal & Glanzer, 1980; Neath, 1993; Whitten, 1978), and the difference between self-paced response and experimenter-paced response has been noted by Poltrock and MacLeod (1977, Experiments 1 and 3) as being a contributing factor in obtaining recency in DFR. In light of these differences, we decided to modify the methodology in Experiment 2 and use 16 seconds of a distractor task consisting of a set of 8 mathematical equations that were presented at an experimenter-controlled rate of 2 seconds per equation.

Secondly, the words presented in Experiment 1 were presented visually on the computer screen and simultaneously spoken via a digitised sound file. By contrast, in many DFR and CDFR studies, the to-be-remembered words were presented silently on the screen (Greene, 1986; Howard & Kahana, 1999; Postman & Phillips, 1965; Raymond, 1969). It is well established that there can be enhanced recency effects for auditory items over visual items in free recall, and these auditory modality advantages occur not just in IFR (e.g., Craik, 1969; Murdock & Walker, 1969), but can also be seen in DFR (Gardiner et al., 1974; Martin and Jones, 1979) and CDFR (Gardiner & Gregg, 1979; Glenberg, 1987; Glenberg & Swanson, 1986; Marks & Crowder, 1997). Therefore a second modification to the methodology used in Experiment 1 was to use visual silent presentation of the words, coupled with an auditory-based distractor task. Specifically, each 16 s of filled distractor activity consisted of 8 mathematical equations that were computer-paced rather than self-paced at a rate of 2s per equation, and the participants' sums in response to the equations were spoken rather than typed.

In summary, Experiment 2 further investigated whether the ISR-like tendencies that are typically obtained in the IFR of short lists remain when a short-term buffer is rendered unavailable. In Experiment 2, participants saw visually-presented lists of between two and twelve words in silence. Again, participants were allocated to one of three groups: IFR, DFR, or CDFR. If the findings in Experiment 1 could be replicated, then we would predict that CDFR and DFR would also exhibit a tendency to initiate recall of short lists with the first list item, suggesting that a short-term buffer cannot be fully responsible for this tendency in IFR. Furthermore, due to the methodological changes adopted, it was hoped that recency at long lists should remain in IFR and CDFR but be eliminated in DFR[3].

## Method

**Participants.** Sixty psychology students from the University of Essex participated in exchange for cash or course credit.

**Materials and apparatus.** The to-be-remembered words were the same as those used in Experiment 1, but they were not accompanied by their corresponding digitised sound files. Furthermore, an expanded set of 113 simple mathematical equations were constructed, which consisted of all the possible additions of three single-digit numbers (e.g., "2 + 1 + 5 = ?") where the correct answer could range from 3 to 10. Using the application Supercard, all stimuli were presented visually in the centre of an Apple Macintosh computer screen. Each participant was provided with a response booklet consisting of 26 response grids, each of which contained two columns and 12 rows. The first column of each grid was narrow and contained the numbers 1-12 in ascending order. The second column was wider to allow room for participants to write down their responses. In order to document answers to the mathematics equations for later inspection, participants' responses were recorded.

**Design.** The design was identical to that used in Experiment 1.

**Procedure.** The procedure was essentially the same as that of Experiment 1 with the exception that each distractor period consisted of solving a series of 8 mathematical equations randomly sampled from the pool of 113 equations described above and were presented silently one at a time in the centre of the computer screen for 2 seconds each. Participants were required to read the equations silently and respond out loud with their answer. After all 8 equations had been shown, participants were presented with the next word in the list (CDFR) or the recall period began (IFR and DFR).

## Results

**Performance on distractor tasks.** The average percentage of equations solved correctly during the mathematics task was 85% in the CDFR condition (with a range of 70-99%) and 87% in the DFR condition (with a range of 70-98%).

**Proportion of words recalled**. The right-hand panels of Figure 1 show the proportion of words recalled for the three tasks using FR scoring (Figure 1B) and SR scoring (Figure 1D). Table 2 summarises the findings of a pair of 3 (task: IFR, DFR and

CDFR) x 6 (list length: 2, 3, 4, 6, 9 and 12) mixed ANOVAs on the proportion of words recalled for both FR and SR scoring.

----------------------------------------

--Table 2 about here--

----------------------------------------

Considering first the data using FR scoring, the significant main effect of list length revealed better performance at shorter list lengths. Specifically, as the list length increased, so the proportion of words recalled decreased monotonically. The significant main effect of task revealed that a greater proportion of words were recalled in the IFR task compared to DFR and CDFR, but the proportions of words recalled in the DFR and the CDFR tasks did not differ significantly. The two-way interaction between task and list length revealed superior IFR performance relative to DFR and CDFR for list lengths: 2, 3, 4, and 6.

Considering next the data using SR scoring, the significant main effect of list length revealed better performance at shorter list lengths. Specifically, as the list length was increased, so the proportion of words recalled in forward serial order decreased monotonically. There was a significant main effect of task, demonstrating that a greater proportion of words were recalled in serial order in IFR, but the degree of forward order recall did not differ significantly between CDFR and DFR. The two-way interaction between task and list length revealed superior serial recall performance in IFR relative to DFR and CDFR for list lengths 2, 3, 4, and 6.

**Analyses of serial position curves of all the data.** Figure 4 shows the serial position curves for each of the three tasks. The left-hand panels show data using FR scoring, with

IFR, DFR, and CDFR shown in Figures 4A, 4C, and 4E, respectively. The right-hand panels show the same data using SR scoring, with IFR, DFR, and CDFR shown in Figures 4B, 4D, and 4F, respectively.

----------------------------------------

--Figure 4 about here--

----------------------------------------

Full statistical analyses of the effects of serial position at each list length and task separately on the proportion of correctly recalled words can be found in Appendix A4 (FR scoring), A5 (SR scoring) and A3 (pairwise comparisons). To summarise the general trends, there were similarities when data were compared across tasks, list lengths and scoring methods.

Considering first the data with FR scoring, the curves changed in similar ways with increasing list length. Performance was close to ceiling for the very short list lengths, but as the list length increased so there were primacy effects at shorter list lengths (more so for IFR); and then reduced primacy at longer list lengths. There were more words recalled in IFR relative to CDFR and DFR (CDFR and DFR did not tend to differ), and this was mainly due to the recency effect being strongest at long list lengths in IFR, but eliminated in DFR and also, surprisingly eliminated in CDFR.

Considering next the data with SR scoring, there was again a similar effect of increasing list length for all tasks. Again, there were more words recalled in IFR relative to CDFR and DFR. However there were no recency effects in any of the tasks using SR scoring which again can be explained if one considers that if participants recalled the recency items, they rarely, if ever, wrote those words down in the final output positions of the response grid, resulting in extended primacy.

**The probability of first response (P[FR]) data.** The right-hand panels of Figure 3 show the proportion of trials in which words from different list positions were recalled first for each list length for IFR, DFR, and CDFR shown in Figure 3B, 3D, and 3F, respectively[4]. As in Experiment 1, the initial output was collapsed into the four categories: ‘SP1’, ‘Last 4’, ‘Other’ or ‘Void / Error’. As can be seen, for all three tasks, recall was most likely to be initiated with the item from the first SP at short list lengths. As the list length increased, this tendency decreased and there was a complimentary increase in tendency to initiate recall with one of the last four words. For IFR and DFR, this resulted in a cross-over: the modal response changed from ‘SP1’ to ‘Last 4’ at list length 9 for IFR and reduced to list length 5 for DFR. There was no cross-over in CDFR: however at list length 6 the two tendencies appeared equiprobable.

Table 2 summarises the findings of a pair of 3 (task: IFR, DFR, and CDFR) x 6 (list length: 2, 3, 4, 6, 9, and 12) mixed ANOVAs conducted on the proportion of trials that were initiated with the word from SP 1 and the proportion of trials that were initiated with one of the last four words.

We will consider first the proportion of trials in which participants initiated recall with the first list item (that is, the P[FR=SP1]). The significant main effects of list length and task revealed a greater tendency to initiate recall with the first list item at shorter list lengths and a greater tendency in IFR (DFR and CDFR did not differ significantly). The two-way interaction between task and list length revealed a superior tendency for IFR at list lengths 2, 3, 4 and 6; at list lengths 9 and 12 there were no significant differences between the tasks.

We will now consider the proportion of trials in which participants initiated recall with one of the last four list items (that is, the P[FR=Last 4]). The significant main effects of list length revealed an increased tendency to initiate recall with one of the last four words with increasing list length. The main effect of task failed to reach significance. The two-way interaction between task and list length revealed that at list lengths 2, this tendency was greatest for CDFR and least for IFR; at list lengths 3 and 4, the tendency was greatest for both CDFR and DFR and least for IFR; at list lengths 9 and 12, the tendency was greatest for IFR and least for both CDFR and DFR; whereas at list length 6, there were no significant differences between the tasks.

## Discussion

In Experiment 2, performance in IFR, DFR, and CDFR was again examined over a range of list lengths, in order to determine whether the tendency to initiate recall of short lists with the first list item, (as is typically found in IFR, Grenfell-Essam & Ward, 2012; Spurgeon et al., 2014; Ward et al., 2010), would additionally be found in DFR and CDFR, conditions in which an explanation based on the direct output of STS would be untenable. Additionally, of interest was whether the recency effects that were unusually present in an attenuated form in DFR in Experiment 1 were now eliminated in Experiment 2.

Experiment 2 replicated the tendency to initiate IFR of a short list of words with the first list item (Ward et al., 2010). Importantly, as in Experiment 1, this same tendency was present, albeit to a reduced extent, in the recall of short lists in DFR and CDFR. There were further similarities between the three tasks. Relative to Experiment 1, there were reduced recency effects in IFR in Experiment 2, perhaps reflecting the change in modality from visual plus auditory presentation to just visual silent presentation. Consistent with prior findings, recency was eliminated in DFR under computer-paced distractor tasks and visual silent presentation of the lists. Surprisingly, the modified methodology also eliminated recency in CDFR. Finally, in all three tasks, the proportion of words recalled decreased with increasing list length.

## General Discussion

This research examined whether an STS explanation could account for why participants often initiated IFR of short lists with the first list item. A number of theorists have proposed that a STS maintains the recency items at test in IFR, following the presentation of long lists (Davelaar et al., 2005; Davelaar et al., 2008, Farrell, 2010; Lehman & Malmberg, 2013; Unsworth & Engle, 2007; Usher et al., 2008). We hypothesised that this same store could potentially also explain why participants tended to initiate IFR of short lists of words with the first list item and proceed in a forward order. To this end, we contrasted performance in IFR over a range of list lengths with both DFR and CDFR (conditions in which list items would be displaced from STS at test).

Our results suggest that the tendency to initiate recall from the beginning of a list cannot be entirely due to a STS, because this tendency remains in CDFR, a variant of free recall in which the contents of any hypothetical STS should be displaced by distractor activity. Although a STS explanation could yet account for the tendency to initiate recall with the first list item in Experiment 1 (because the contents of STS may not have been fully displaced – as evidenced by only partially attenuated recency in DFR), such an explanation could not account for the data from CDFR in Experiment 2, where recency in DFR was entirely eliminated.

These findings strongly suggest that the tendency is at the very least not entirely attributable to STS. Rather, it would seem that the tendency to initiate short lists of items with the first list item might be a more general property of episodic memory. Because the tendency to initiate recall of short lists with the first list item was greatest in IFR, dual-store theorists could reasonably assume that STS may augment the tendency under those situations in which its use is tenable. However, there must be some question over whether it is parsimonious to propose separate short-term and long-term memory mechanisms for the three properties of recency effects, contiguity effects, and now first response effects.

Nevertheless, our data continue to cause difficulties for current unitary models of memory. Although these models can explain recency effects at all time scales (e.g., Brown et al., 2007; Howard & Kahana, 2002), and some can naturally account for temporal contiguity effects (e.g., Howard & Kahana, 2002), these accounts appear to require an additional assumption or mechanism to explain the far greater accessibility of the first list item relative to the last item in short lists. It should therefore be acknowledged that unitary models are not necessarily more parsimonious than a dual store model when it comes to explaining first response effects.

Quite what drives the tendency to initiate recall of short lists with the first list item remains uncertain. It is possible that the tendency reflects the privileged access to the start of the list or the start of the current group (Farrell, 2012) affording a special status to the first list item. Perhaps the first item benefits from the retrieval of a 'start-of-list' marker (Davelaar et al., 2005), a 'get ready warning' signal used prior to the presentation of the first list item (Laming 1999, 2010), greater positional certainty (Henson, 1998) or an association with an internal contextual state which may allow selective access to the first item (Metcalfe & Murdock, 1981). An alternative possibility is that there is preferential access to the first item in free recall, due to the increased attention that is attributed to the first item during presentation (Raijmaakers & Shiffrin, 1981), or the increased novelty of early list items relative to later list items (Brown, Preece & Hulme, 2000; Farrell & Lewandowsky, 2002; Lewandowsky & Farrell, 2008). This special status of the first list item, whatever form it may take, becomes less accessible as list length increases or time passes.

However, a number of plausible alternatives have been ruled out. We have recently shown that the tendency is neither underpinned by rehearsal (Grenfell-Essam *et al*., 2013), nor requires phonological coding (Spurgeon *et al.*, 2014). Additional research from our laboratory clearly shows the same effects of list length and output order occur in the IFR of non-verbal stimuli (the IFR of different length sequences of visuo-spatial dots, Cortis, Dent, Kennett, & Ward, under review), a finding suggesting that the finding is not limited to the verbal domain. Thus, the current findings suggest that this tendency reflects a general property of episodic memory that can be observed over a range of different timescales.

Finally, our data from DFR and CDFR suggest that there is a natural tendency to initiate recall of short lists with the first list item and continue in forwards ordered recall, even following filled periods of distractor activity that are presented at the end of the list, or after each and every list item. These findings raise an interesting parallel to the serial recall literature. Ward *et al*. (2010) argued for greater theoretical integration between IFR and ISR based largely on the "ISR-like" tendencies observed in the IFR of short lists. Although highly speculative, our current data suggest that there might be related similarities between DFR and delayed serial recall (that is, the Brown-Peterson task), and CDFR and continuous distractor serial recall (that is, the working memory span task), potentially encouraging yet wider integration of memory tasks. Although one can confidently predict that the to-be-remembered words in conventional STS buffers should be displaced by distractor activity, it is possible that the immediate memory mechanisms postulated in more complex models of working memory span (e.g., Barrouillet, Bernardin & Camos, 2004; Barrouillet, Gavens, Vergauwe, Gaillard & Camos, 2009; Oberauer, Lewandowsky, Farrell, Jarrold and Greaves, 2012; Unsworth & Engle, 2007) would be better suited to juggle the demands of the distractors at the same time as maintaining the words in memory.

## References

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: advances in research and theory, Vol. 2.* (pp. 89-195). New York: Academic Press.

Atkinson, R. C., & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American, 225,* 82-90.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.) *Recent advances in learning and motivation, Vol. 8.* (pp. 47-90). London: Academic Press.

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

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

Beaman, C. P., & Morton, J. (2000). The separate but related origins of the recency and the modality effect in free recall. *Cognition,* 77, B59-B65.

Bhatarah, P., Ward, G., Smith, J., & Hayes, L. (2009). Examining the relationship between free recall and immediate serial recall: Similar patterns of rehearsal and similar effects of word length, presentation rate, and articulatory suppression. *Memory and Cognition, 37,* 689-713.

Bhatarah, P., Ward, G., & Tan, L. (2006). Examining the relationship between immediate serial recall and free recall: The effect of concurrent task performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 215-229.

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

Brodie, D. A., & Murdock, B. B., Jr. (1977). Effect of presentation time on nominal and functional serial-position curves of free recall. *Journal of Verbal Learning and Verbal Behavior, 16,* 185–200.

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

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

Cortis, C., Dent, K., Kennett, S., Ward, G. (under review). First things first: Similar list length and output order effects for verbal and non-verbal stimuli. *Journal of Experimental Psychology: Learning, Memory & Cognition.*

Craik, F. I. M. (1969). Modality effects in short-term memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 658-664.

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., Usher, M., Haarmann, H. J., & Goshen-Gottstein, Y. (2008). Through TCM, STM shines bright. *Psychological Review, 115,* 1116-1118.

Farrell, S. (2010). Dissociating conditional recency in immediate and delayed free recall: A challenge for unitary models of recency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 324-347.

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

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

Friendly, M., Franklin, P. E., Hoffman, D., & Rubin, D. C. (1982). Norms for the Toronto Word Pool: Norms for imagery, concreteness, orthographic variables and grammatical usage for 1,080 words. *Behavior Research Methods & Instrumentation*, *14,* 375-399.

Hurlstone, M. J., Hitch, G. J., & Baddeley, A. D. (2014). Memory for serial order across domains: An overview of the literature and directions for future research. *Psychological Bulletin, 140*, 339-373.

Gardiner, J. M., & Gregg, V. H. (1979). When auditory memory is not overwritten. *Journal of Verbal Learning and Verbal Behavior, 18*, 705-719.

Gardiner, J. M., Thompson, C. P., & Maskarinec, S. (1974). Negative recency in free recall. *Journal of Experimental Psychology*, *103,* 71-78.

Glanzer, M. (1972). Storage mechanisms in recall. In G. H. Bower, (Ed.), *The psychology of learning and motivation: Advances in research and theory*. (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.

Glenberg, A. M. (1987). Temporal context and recency. In D. S. Gorfein & R. R. Hoffman, (Eds.), *Memory and learning: The Ebbinghaus Centennial Conference* (pp. 173-190). Lawrence Erlbaum Associates, Inc. Hillsdale, New Jersey.

Glenberg, A. M., Bradley, M. M., Kraus, T. A., & Ranzaglia, G. J. (1983). Studies of the long-term recency effect: Support for a contextually guided retrieval theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 413-418.

Glenberg, A. M., Bradley, M. M., Stevenson, J. A., Kraus, T. A., Tkachuk, M. J., & Gretz, A. L. (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. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 3-15.

Greene, R. L. (1986). A common basis for recency effects in immediate and delayed recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*, 413–418.

Greenhouse, S.W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24,* 95-112.

Grenfell-Essam, R., & Ward, G. (2012). Examining the relationship between free recall and immediate serial recall: The role of list length, strategy use, and test expectancy. *Journal of Memory and Language, 67,* 106-148.

Grenfell-Essam, R., Ward, G., & Tan, L. (2013). The role of rehearsal on the output order of immediate free recall of short and long lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39,* 317-347.

Henson, R. N. A. (1998). Short-term memory for serial order: The start-end model of serial recall. *Cognitive Psychology, 36,* 73-137.

Hogan, R. M. (1975). Interitem encoding and directed search in free recall. *Memory & Cognition, 3,* 197–209.

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.

Howard, M. W., Youker, T. E., & Venkatadass, V. S. (2008). The persistence of memory: Contiguity effects across hundreds of seconds. *Psychonomic Bulletin & Review, 15,* 58-63.

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

Kahana, M. J. (2012a). Auditory Toronto word pool. Available : http://memory.psych.upenn.edu/WordPools (11th September, 2012).

Kahana, M. J. (2012b). *Foundations of Human Memory.* New York: Oxford University Press.

Koppenaal, L., & Glanzer, M. (1990). An examination of the continuous distractor task and the long-term recency effect. *Memory and Cognition, 18,* 183-195.

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.

Laming, D. (2008). An improved algorithm for predicting free recalls. *Cognitive Psychology, 57*, 179–219.

Laming, D. (2009). Failure to recall. *Psychological Review, 116*, 157–186.

Laming, D. (2010). Serial position curves in free recall. *Psychological Review, 117*, 93-113.

Lehman, M. & Malmberg, K. J. (2013). A buffer model of encoding and temporal correlations in retrieval. *Psychological Review*, *120*, 155-189.

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. *The Psychology of Learning and Motivation, 49*, 1-48.

Marks, A. R., & Crowder, R. G. (1997). Temporal distinctiveness and modality. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 164-180.

Martin, M., & Jones, G. V. (1979). Modality dependency of loss of recency in free recall. *Psychological Research, 40*, 273-289.

Metcalfe, J., & Murdock, B. B. (1981). An encoding and retrieval model of single-trial free recall. *Journal of Verbal Learning and Verbal Behavior, 20*, 161–189.

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

Moreton, B. J., & Ward, G. (2010). Time scale similarity and long-term memory for autobiographical events. *Psychonomic Bulletin & Review*, *17,* 510-515.

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

Murdock, B. B., Jr., & Walker, K. D. (1969). Modality effects in free recall. *Journal of Verbal Learning and Verbal Behavior, 8,* 665-676.

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. (1993). Contextual and distinctive processes and the serial position function. *Journal of Memory and Language, 32*, 820-840.

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

Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C., & Greaves, M. (2012). Modeling working memory: An interference model of complex span. *Psychonomic Bulletin & Review, 19,* 779-819.

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.

Poltrock, S. E., & MacLeod, C. M. (1977). Primacy and recency in the continuous distractor paradigm. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 560–571.

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

Postman, L., & Phillips, L. W. (1965). Short-term temporal change in free recall. *Quarterly Journal of Experimental Psychology, 56,* 413-419.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134.

Raymond, B. J. (1969). Short-term storage and long-term storage in free recall. *Journal of Verbal Learning and Verbal Behavior, 8,* 567-574.

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

Sederberg, P. B., Howard, M. W., & Kahana, M. J. (2008). A context-based theory of recency and contiguity in free recall. *Psychological Review, 115,* 893-912.

Spurgeon, J. Ward, G., & Matthews, W.J. (2014). Examining the relationship between immediate serial recall and immediate free recall: Common effects of Phonological Loop variables, but only limited evidence for the Phonological Loop. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Advance online publication.

Sumby, W.H. (1963). Word frequency and serial position effects. *Journal of Verbal Learning and Verbal Behavior*, *1,* 443-450.

Surprenant, A. M., & Neath, I. (2009). The nine lives of short-term memory. In A. Thorn & M. Page (Eds.). *Interactions between short-term and long-term memory in the verbal domain* (pp. 16-43). Hove, East Sussex: Psychology Press.

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

Thorn, A. S. C. & Page, M. P. A. (2008). *Interactions between short-term and long-term memory in the verbal domain*. Hove, UK: Psychology Press.

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

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114,* 104-132.

Usher, M., Davelaar, E. J., Haarmann, H. J., & Goshen-Gottstein, Y. (2008). Short-term memory after all: Comment on Sederberg, Howard, and Kahana (2008). *Psychological Review, 115,* 1108-1118.

Ward, G. (2002). A recency-based account of the list length effect in free recall. *Memory & Cognition, 30,* 885–892.

Ward, G. & Tan, L. (2004). The effect of the length of to-be-remembered lists and intervening lists on free recall: A re-examination using overt rehearsal. *Journal of Experimental Psychology: Learning, Memory & Cognition, 30,* 1196-1210.

Ward, G., Tan, L., & Grenfell-Essam, R. (2010). Examining the relationship between free recall and immediate serial recall: The effects of list length and output order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 1207-1241.

Ward, G., Woodward, G., Stevens, A., & Stinson, C. (2003). Using overt rehearsals to explain word frequency effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 186–210.

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

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

Whitten, W. B. (1978). Output interference and long-term serial position effects. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 685-692.

# Footnotes

[1] Note the analyses of the resultant serial position curves conditionalised by trials in which recall was initiated with the word from SP1 or with one of the last four words can be obtained from the first author. When P(FR=SP1), all three tasks showed extended primacy (SR and FR scoring) with recency (FR scoring). When P(FR=Last 4), all three tasks showed extended recency and virtually no primacy (FR scoring).

[2] Note that for list lengths 2-4, the Last 4 category excludes trials starting with the word from serial position 1.

[3] We first employed these modifications in a preliminary experiment to Experiment 2, which, following Greene (1986), examined DFR at a fixed list length of 8 words. Words were presented visually and silently, and recall was written. Each distractor period consisted of 8 equations presented visually and silently for 2 seconds each, and participants made their responses to the equations aloud. Eighteen participants each performed 24 trials. This preliminary experiment showed highly significant effects of serial position, $F(7,119) = 6.60$, $MSE = .013$, $p < .001$, $\eta^2_p = .280$ (FR scoring); follow-up tests confirmed there was a significant primacy effect but an absence of recency {the means for serial positions 1-8 were: .45, .32, .33, .26, .28, .22, .28, .27, respectively].

[4] Note the analyses of the resultant serial position curves conditionalised by trials in which recall was initiated with the word from SP1 or with one of the last four words can be obtained from the first author. When P(FR=SP1), all three tasks showed extended

primacy with no recency (SR and FR scoring). When P(FR=Last 4), all three tasks showed extended recency and virtually no primacy (FR scoring).

# Table Captions

*Table 1.* Summary of the ANOVA tables from Experiment 1. The analyses were conducted upon the proportion of correctly recalled words using FR scoring and SR scoring, and the Probability of First Response (P[FR]) data (LL = list length). Note, sphericity tests were performed which confirmed no violations of assumptions.

*Table 2.* Summary of the ANOVA tables from Experiment 2. The analyses were conducted upon the proportion of correctly recalled words using FR scoring and SR scoring, and the Probability of First Response (P[FR]) data (LL = list length). Note, sphericity tests were performed which confirmed no violations of assumptions.

**Table 1.**

| | *df* | *MSE* | *F* | $\eta^2_p$ | *p* |
|---|---|---|---|---|---|
| *Proportion correct (FR scoring)* | | | | | |
| Task | 2, 57 | .063 | 6.12 | .177 | =.004 |
| LL | 5, 285 | .012 | 261.74 | .821 | <.001 |
| Task x LL | 10, 285 | .012 | 5.38 | .159 | <.001 |
| *Proportion correct (SR scoring)* | | | | | |
| Task | 2, 57 | .111 | 11.52 | .288 | <.001 |
| LL | 5, 285 | .026 | 256.04 | .818 | <.001 |
| Task x LL | 10, 285 | .026 | 9.36 | .247 | <.001 |
| *P(FR=SP1)* | | | | | |
| Task | 2, 57 | .229 | 9.34 | .247 | <.001 |
| LL | 5, 285 | .049 | 89.57 | .611 | <.001 |
| Task x LL | 10, 285 | .049 | 5.75 | .168 | <.001 |
| *P(FR=Last 4)* | | | | | |
| Task | 2, 57 | .248 | 3.92 | .121 | =.025 |
| LL | 5, 285 | .067 | 26.63 | .318 | <.001 |
| Task x LL | 10, 285 | .067 | 6.38 | .183 | <.001 |

**Table 2.**

| | *df* | *MSE* | *F* | $\eta^2_p$ | *p* |
|---|---|---|---|---|---|
| *Proportion correct (FR scoring)* | | | | | |
| Task | 2, 57 | .124 | 14.73 | .341 | <.001 |
| LL | 5, 285 | .013 | 176.22 | .756 | <.001 |
| Task x LL | 10, 285 | .013 | 8.10 | .221 | <.001 |
| *Proportion correct (SR scoring)* | | | | | |
| Task | 2, 57 | .122 | 19.82 | .410 | <.001 |
| LL | 5, 285 | .019 | 230.05 | .801 | <.001 |
| Task x LL | 10, 285 | .019 | 22.83 | .445 | <.001 |
| *P(FR=SP1)* | | | | | |
| Task | 2, 57 | .210 | 11.15 | .281 | <.001 |
| LL | 5, 285 | .046 | 56.19 | .496 | <.001 |
| Task x LL | 10, 285 | .046 | 6.66 | .189 | <.001 |
| *P(FR=Last 4)* | | | | | |
| Task | 2, 57 | .128 | 2.18 | .071 | =.123 |
| LL | 5, 285 | .046 | 16.33 | .223 | <.001 |
| Task x LL | 10, 285 | .046 | 11.98 | .296 | <.001 |

# Figure Captions

*Figure 1.* Data from Experiments 1 and 2 showing the mean proportion of words recalled from lists of 2 to 12 words presented for IFR, DFR and CDFR. The left-hand panels show the mean proportion of words recalled from Experiment 1 using FR scoring (Figure 1A) and SR scoring (Figure 1C), and the right-hand panels show the mean proportion of words recalled from Experiment 2 using FR scoring (Figure 1B) and SR scoring (Figure 1D).

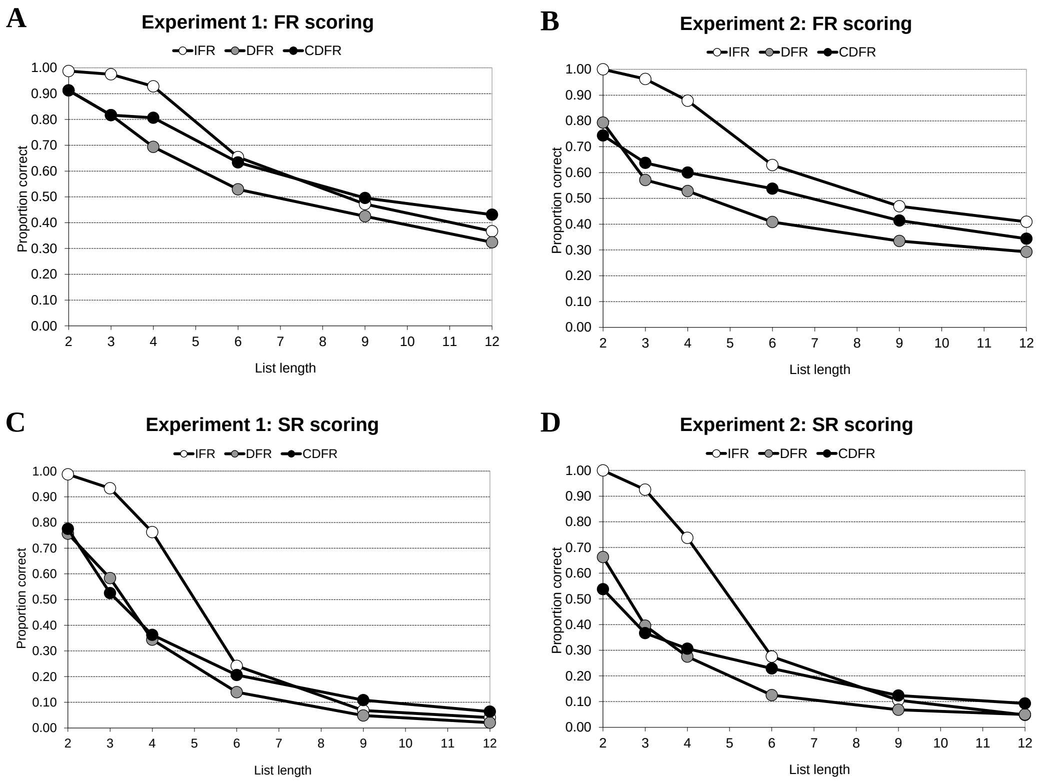
*Figure 2.* Data from Experiment 1 showing the serial position curves from lists of 2 to 12 words. The left-hand panels show the serial position curves using FR scoring: for IFR (Figure 2A), DFR (Figure 2C) and CDFR (Figure 2E); the right-hand panels show the serial position curves using SR scoring: for IFR (Figure 2B), DFR (Figure 2D) and CDFR (Figure 2F).

*Figure 3.* The Probability of First Response (P[FR]) data from Experiments 1 and 2. These plots show the proportion of trials in which recall initiated with the word from SP 1 in the list, one of the last four words on the list, or one of the other words in the list. On a small minority of trials, participants began recall with an error. Each task is plotted separately by increasing list length. The left-hand panels show the data from Experiment 1 for IFR (Figure 3A), DFR (Figure 3C), and CDFR (Figure 3E); the right-hand panels show the data from Experiment 2 for IFR (Figure 3B), DFR (Figure 3D), and CDFR (Figure 3F).

*Figure 4.* Data from Experiment 2 showing the serial position curves from lists of 2 to 12 words. The left-hand panels show the serial position curves using FR scoring: for IFR (Figure 4A), DFR (Figure 4C) and CDFR (Figure 4E); the right-hand

panels show the serial position curves using SR scoring: for IFR (Figure 4B), DFR (Figure 4D) and CDFR (Figure 4F).

**Figure 1.**


A
Experiment 1: FR scoring
IFR DFR CDFR
Proportion correct
List length
B
Experiment 2: FR scoring
IFR DFR CDFR
Proportion correct
List length
C
Experiment 1: SR scoring
IFR DFR CDFR
Proportion correct
List length
D
Experiment 2: SR scoring
IFR DFR CDFR
Proportion correct
List length

# Figure 2.

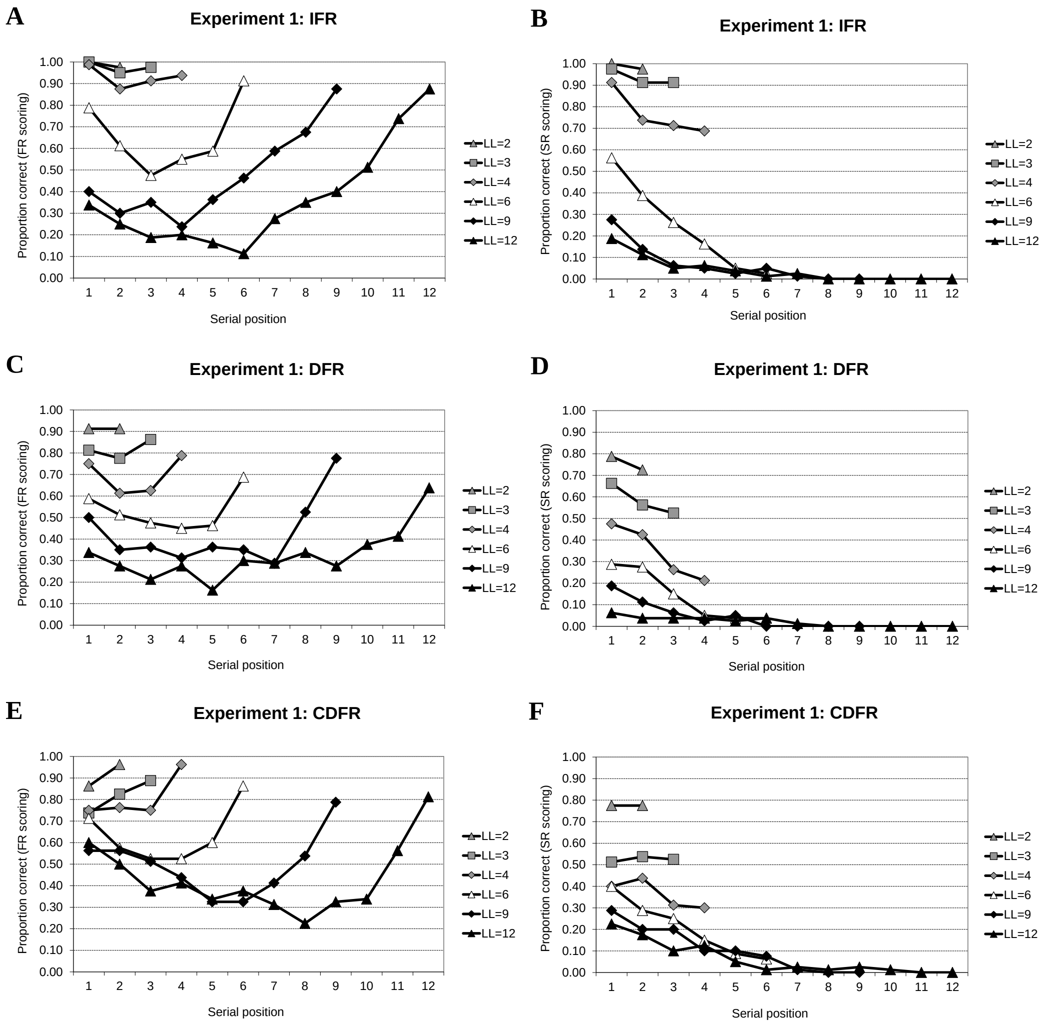

A
Experiment 1: IFR
Proportion correct (FR scoring)
Serial position
LL=2
LL=3
LL=4
LL=6
LL=9
LL=12
B
Experiment 1: IFR
Proportion correct (SR scoring)
Serial position
C
Experiment 1: DFR
Proportion correct (FR scoring)
Serial position
D
Experiment 1: DFR
Proportion correct (SR scoring)
Serial position
E
Experiment 1: CDFR
Proportion correct (FR scoring)
Serial position
F
Experiment 1: CDFR
Proportion correct (SR scoring)
Serial position

**Figure 3.**

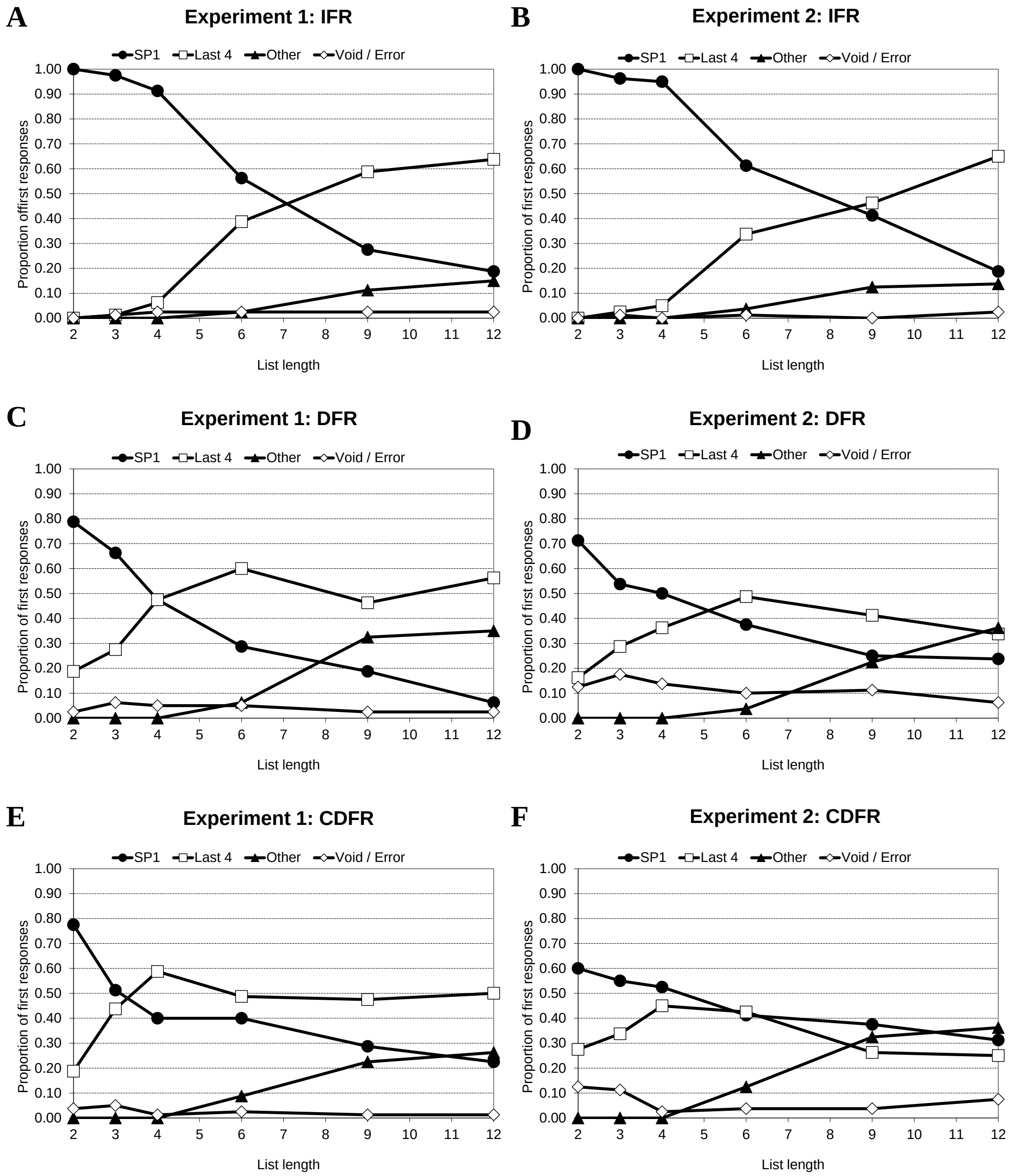
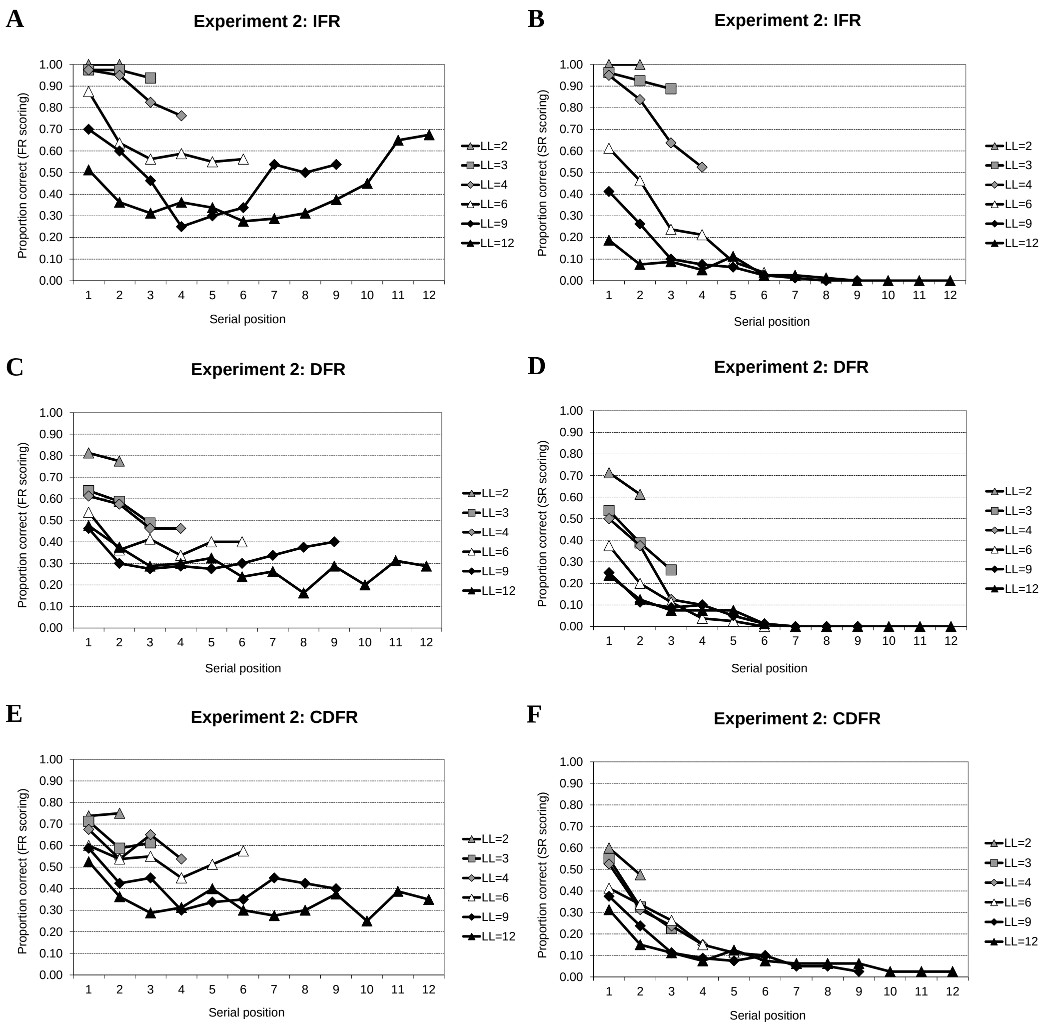

A
Experiment 1: IFR
SP1
Last 4
Other
Void / Error
Proportion offirst responses
List length
B
Experiment 2: IFR
Proportion of first responses
List length
C
Experiment 1: DFR
Proportion of first responses
List length
D
Experiment 2: DFR
Proportion of first responses
List length
E
Experiment 1: CDFR
Proportion of first responses
List length
F
Experiment 2: CDFR
Proportion of first responses
List length

**Figure 4.**


A
Experiment 2: IFR
Proportion correct (FR scoring)
Serial position
LL=2
LL=3
LL=4
LL=6
LL=9
LL=12
B
Experiment 2: IFR
Proportion correct (SR scoring)
Serial position
LL=2
LL=3
LL=4
LL=6
LL=9
LL=12
C
Experiment 2: DFR
Proportion correct (FR scoring)
Serial position
LL=2
LL=3
LL=4
LL=6
LL=9
LL=12
D
Experiment 2: DFR
Proportion correct (SR scoring)
Serial position
LL=2
LL=3
LL=4
LL=6
LL=9
LL=12
E
Experiment 2: CDFR
Proportion correct (FR scoring)
Serial position
LL=2
LL=3
LL=4
LL=6
LL=9
LL=12
F
Experiment 2: CDFR
Proportion correct (SR scoring)
Serial position
LL=2
LL=3
LL=4
LL=6
LL=9
LL=12

## Appendix Captions

*Appendix A1*. Analyses of the serial position curves from Experiment 1 using FR scoring. At each list length (LL) and task, the data were subjected to a *n* (serial position (SP): 1,…,*n*) single factor within-subjects ANOVA, where *n* is the list length. Significant main effects and interactions are shown in **bold**. Note, sphericity tests were performed, and for those effects marked with an asterix, the results were corrected using the Greenhouse-Geisser (1959) correction.

*Appendix A2*. Analyses of the serial position curves from Experiment 1 using SR scoring. At each list length (LL) and task, the data were subjected to a *n* (serial position (SP): 1,…,*n*) single factor within-subjects ANOVA, where *n* is the list length. Significant main effects and interactions are shown in **bold**. Note, sphericity tests were performed, and for those effects marked with an asterix, the results were corrected using the Greenhouse-Geisser (1959) correction.

*Appendix A3*. Summary of the extent of the primacy and recency effects for those significant effects of serial position (SP) observed in Appendices A1, A2, A4 and A5. To define the extent of primacy effects, the SP with the lowest performance was identified. Primacy effects were then determined by observing the number of significant pairwise comparison steps between SP1 and that lowest SP. The extent of primacy is reproduced by the number of +'s, such that '+ +' represents a significant decrease in recall between SP1 and SP*a*, and a further significant decrease in recall between SP*a* and SP*b*, where *a* is earlier in the list than *b*. Recency effects were calculated in the same manner by observing the number of significant pairwise comparison steps between the lowest SP and SP*n* (where *n* is the last list item).

*Appendix A4.* Analyses of the serial position curves from Experiment 2 using FR scoring. At each list length (LL) and task, the data were subjected to a *n* (serial position (SP): 1,…,*n*) single factor within-subjects ANOVA, where *n* is the list length. Significant main effects and interactions are shown in **bold**. Note, sphericity tests were performed, and for those effects marked with an asterix, the results were corrected using the Greenhouse-Geisser (1959) correction.

*Appendix A5.* Analyses of the serial position curves from Experiment 2 using SR scoring. At each list length (LL) and task, the data were subjected to a *n* (serial position (SP): 1,…,*n*) single factor within-subjects ANOVA, where *n* is the list length. Significant main effects and interactions are shown in **bold**. Note, sphericity tests were performed, and for those effects marked with an asterix, the results were corrected using the Greenhouse-Geisser (1959) correction.

## Appendix A1.

| | IFR | DFR | CDFR |
|---|---|---|---|
| LL 2 | $F(1,19) = 2.11$, $MSE = .003$, $p = .163$, $\eta^2_p = .100$ | $F(1,19) = 0.00$, $MSE = .020$, $p = 1$, $\eta^2_p < .001$ | **$F(1,19) = 5.63$, $MSE = .018$, $p = .028$, $\eta^2_p = .229$** |
| LL 3 | * $F(1,25) = 1.54$, $MSE = .012$, $p = .232$, $\eta^2_p = .075$ | $F(2,38) = 1.66$, $MSE = .023$, $p = .203$, $\eta^2_p = .080$ | $F(2,38) = 2.38$, $MSE = .048$, $p = .106$, $\eta^2_p = .111$ |
| LL 4 | **$F(3,57) = 3.83$, $MSE = .012$, $p = .014$, $\eta^2_p = .168$** | **$F(3,57) = 3.63$, $MSE = .043$, $p = .018$, $\eta^2_p = .160$** | **$F(3,57) = 5.83$, $MSE = .037$, $p = .002$, $\eta^2_p = .235$** |
| LL 6 | **$F(5,95) = 7.97$, $MSE = .067$, $p < .001$, $\eta^2_p = .296$** | **$F(5,95) = 2.67$, $MSE = .063$, $p = .026$, $\eta^2_p = .123$** | **$F(5,95) = 6.44$, $MSE = .054$, $p < .001$, $\eta^2_p = .253$** |
| LL 9 | **$F(8,152) = 14.25$, $MSE = .059$, $p < .001$, $\eta^2_p = .429$** | * **$F(5,90) = 8.08$, $MSE = .098$, $p < .001$, $\eta^2_p = .298$** | * **$F(4,72) = 5.77$, $MSE = .150$, $p = .001$, $\eta^2_p = .233$** |
| LL 12 | * **$F(5,99) = 21.43$, $MSE = .109$, $p < .001$, $\eta^2_p = .530$** | * **$F(6,110) = 5.11$, $MSE = .106$, $p < .001$, $\eta^2_p = .212$** | **$F(11,209) = 8.40$, $MSE = .062$, $p < .001$, $\eta^2_p = .307$** |

## Appendix A2.

| | IFR | DFR | CDFR |
|---|---|---|---|
| LL 2 | $F(1,19) = 2.11$, $MSE = .003$, $p = .163$, $\eta^2_p = .100$ | $F(1,19) = 4.13$, $MSE = .009$, $p = .056$, $\eta^2_p = .179$ | $F(1,19) = 0.00$, $MSE = .003$, $p = 1$, $\eta^2_p < .001$ |
| LL 3 | * $F(1,24) = 2.02$, $MSE = .020$, $p = .146$, $\eta^2_p = .096$ | * $F(1,27) = 3.66$, $MSE = .039$, $p = .053$, $\eta^2_p = .162$ | * $F(1,28) = 0.14$, $MSE = .030$, $p = .802$, $\eta^2_p = .007$ |
| LL 4 | * **$F(2,35) = 8.84$, $MSE = .038$, $p = .001$, $\eta^2_p = .317$** | * **$F(2,38) = 14.67$, $MSE = .033$, $p < .001$, $\eta^2_p = .436$** | **$F(3,57) = 3.95$, $MSE = .023$, $p = .013$, $\eta^2_p = .172$** |
| LL 6 | * **$F(2,46) = 21.16$, $MSE = .084$, $p < .001$, $\eta^2_p = .527$** | * **$F(3,57) = 15.57$, $MSE = .029$, $p < .001$, $\eta^2_p = .450$** | * **$F(2,40) = 8.91$, $MSE = .090$, $p = .001$, $\eta^2_p = .319$** |
| LL 9 | **$F(8,152) = 6.71$, $MSE = .023$, $p < .001$, $\eta^2_p = .261$** | **$F(8,152) = 8.16$, $MSE = .010$, $p < .001$, $\eta^2_p = .300$** | **$F(8,152) = 8.72$, $MSE = .024$, $p < .001$, $\eta^2_p = .315$** |
| LL 12 | **$F(11,209) = 6.12$, $MSE = .011$, $p < .001$, $\eta^2_p = .244$** | **$F(11,209) = 2.06$, $MSE = .004$, $p = .024$, $\eta^2_p = .098$** | **$F(11,209) = 8.62$, $MSE = .013$, $p < .001$, $\eta^2_p = .312$** |

## Appendix A3.

| | | Experiment 1 | | | | Experiment 2 | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | FR scoring | | SR scoring | | FR scoring | | SR scoring | |
| | | Primacy | Recency | Primacy | Recency | Primacy | Recency | Primacy | Recency |
| LL 2 | IFR | | | | | | | | |
| | DFR | | | | | | | + | |
| | CDFR | | + | | | | | + | |
| LL 3 | IFR | | | | | | | + | |
| | DFR | | | | | | | + + | |
| | CDFR | | | | | | | + + | |
| LL 4 | IFR | + | | + | | + | | + + | |
| | DFR | + | + | + | | | | + + | |
| | CDFR | | + + | + | | | | + + | |
| LL 6 | IFR | + + | + | + + + | | + | | + + + | |
| | DFR | | + | + + | | | | + + | |
| | CDFR | + | + | + + | | | | + + | |
| LL 9 | IFR | | + + + | + + | | + + | + | + + + | |
| | DFR | + | + + | + + | | | | + + | |
| | CDFR | + | + + | + + | | + | | + + + | |
| LL 12 | IFR | + | + + + + | + + | | + | + + | + + | |
| | DFR | + | + + | + | | + + | | + + | |
| | CDFR | + + | + + | + + | | + | | + + | |

## Appendix A4.

| | IFR | DFR | CDFR |
|---|---|---|---|
| LL 2 | $F(1,19) = 0.00$, $MSE = .000$, $p = 1$, $\eta^2_p < .001$ | $F(1,19) = 1.31$, $MSE = .011$, $p = .267$, $\eta^2_p = .064$ | $F(1,19) = 0.04$, $MSE = .038$, $p = .841$, $\eta^2_p = .002$ |
| LL 3 | $F(2,38) = 1.88$, $MSE = .005$, $p = .167$, $\eta^2_p = .090$ | $F(2,38) = 2.03$, $MSE = .057$, $p = .145$, $\eta^2_p = .097$ | $F(2,38) = 1.71$, $MSE = .051$, $p = .195$, $\eta^2_p = .082$ |
| LL 4 | ***$F(2,41) = 6.82$, $MSE = .042$, $p = .001$, $\eta^2_p = .264$** | $F(3,57) = 2.04$, $MSE = .059$, $p = .119$, $\eta^2_p = .097$ | $F(3,57) = 2.02$, $MSE = .053$, $p = .121$, $\eta^2_p = .096$ |
| LL 6 | ***$F(3,53) = 3.79$, $MSE = .147$, $p = .018$, $\eta^2_p = .166$** | $F(5,95) = 1.96$, $MSE = .049$, $p = .092$, $\eta^2_p = .093$ | $F(5,95) = 1.03$, $MSE = .053$, $p = .405$, $\eta^2_p = .051$ |
| LL 9 | ***$F(4,77) = 6.24$, $MSE = .139$, $p < .001$, $\eta^2_p = .247$** | *$F(4,82) = 1.37$, $MSE = .115$, $p = .216$, $\eta^2_p = .067$ | ***$F(5,91) = 2.28$, $MSE = .061$, $p = .025$, $\eta^2_p = .107$** |
| LL 12 | **$F(11,209) = 5.87$, $MSE = .063$, $p < .001$, $\eta^2_p = .236$** | **$F(11,209) = 2.31$, $MSE = .056$, $p = .011$, $\eta^2_p = .108$** | **$F(11,209) = 2.30$, $MSE = .048$, $p = .011$, $\eta^2_p = .108$** |

## Appendix A5.

| | IFR | DFR | CDFR |
|---|---|---|---|
| LL 2 | $F(1,19) = 0.00$, $MSE = .000$, $p = 1$, $\eta^2_p < .001$ | **$F(1,19) = 8.94$, $MSE = .011$, $p = .008$, $\eta^2_p = .320$** | **$F(1,19) = 7.31$, $MSE = .021$, $p = .014$, $\eta^2_p = .278$** |
| LL 3 | ***$F(1,27) = 3.35$, $MSE = .008$, $p = .046$, $\eta^2_p = .150$** | **$F(2,38) = 13.94$, $MSE = .027$, $p < .001$, $\eta^2_p = .423$** | ***$F(1,28) = 20.71$, $MSE = .036$, $p < .001$, $\eta^2_p = .522$** |
| LL 4 | ***$F(2,45) = 22.97$, $MSE = .040$, $p < .001$, $\eta^2_p = .547$** | ***$F(2,41) = 18.39$, $MSE = .057$, $p < .001$, $\eta^2_p = .492$** | **$F(3,57) = 16.82$, $MSE = .031$, $p < .001$, $\eta^2_p = .470$** |
| LL 6 | ***$F(3,49) = 27.88$, $MSE = .069$, $p < .001$, $\eta^2_p = .595$** | ***$F(2,39) = 20.25$, $MSE = .020$, $p < .001$, $\eta^2_p = .516$** | ***$F(2,47) = 10.49$, $MSE = .064$, $p < .001$, $\eta^2_p = .356$** |
| LL 9 | **$F(8,152) = 14.32$, $MSE = .028$, $p < .001$, $\eta^2_p = .430$** | **$F(8,152) = 8.24$, $MSE = .016$, $p < .001$, $\eta^2_p = .303$** | **$F(8,152) = 12.49$, $MSE = .020$, $p < .001$, $\eta^2_p = .397$** |
| LL 12 | **$F(11,209) = 5.29$, $MSE = .013$, $p < .001$, $\eta^2_p = .218$** | **$F(11,209) = 8.77$, $MSE = .012$, $p < .001$, $\eta^2_p = .316$** | **$F(11,209) = 10.41$, $MSE = .012$, $p < .001$, $\eta^2_p = .354$** |