Journal of Experimental Psychology:
Learning, Memory, and Cognition
2010, Vol. 36, No. 5, 1207–1241

0278-7393/10/$12.00 DOI: 10.1037/a0020122

Examining the Relationship Between Free Recall and Immediate Serial Recall: The Effects of List Length and Output Order

Geoff Ward
University of Essex

Lydia Tan
City University

Rachel Grenfell-Essam
University of Essex

In 4 experiments, participants were presented with lists of between 1 and 15 words for tests of immediate memory. For all tasks, participants tended to initiate recall with the first word on the list for short lists. As the list length was increased, so there was a decreased tendency to start with the first list item; and, when free to do so, participants showed an increased tendency to start with one of the last 4 list items. In all tasks, the start position strongly influenced the shape of the resultant serial position curves: When recall started at Serial Position 1, elevated recall of early list items was observed; when recall started toward the end of the list, there were extended recency effects. These results occurred under immediate free recall (IFR) and different variants of immediate serial recall (ISR) and reconstruction of order (RoO) tasks. We argue that these findings have implications for the relationship between IFR and ISR and between rehearsal and recall.

Keywords: free recall, serial recall, short-term memory, output order, list length

The overall aim of this research is to promote greater theoretical integration between two highly important and widely used tests of immediate memory: immediate serial recall (ISR) and immediate free recall (IFR). The main claim of the article is that greater theoretical integration between these tasks could be achieved, if only researchers understood fully the effects of increasing list length on both tasks. To this end, we report the data from four experiments looking at the effects of increasing list length on the output order and serial position curves for the IFR task, the ISR task, and variants of the ISR task.

At first glance, one might think that there should be no need to promote greater theoretical integration between ISR and IFR. On the face of it, the methodologies of the two immediate memory tests are remarkably similar: In both tasks, participants are presented with a sequence of items and at the end of the list they must try to recall as many items as possible, in either the same order as that presented (ISR) or in any order (IFR). In addition, both tasks share a common theoretical heritage. Both tasks have provided empirical evidence taken as key "signature findings" supporting the establishment of a short-term memory store (STS) of limited capacity: In ISR, the memory span limitation has been taken to reflect the limited capacity of verbal STS (whether this be measured in items, chunks, or time; e.g., Baddeley, 1986; Miller, 1956); in IFR, the advantage in recall of the last items, known as the *recency effect*, has also been taken as evidence for the direct output of items at test from a short-term buffer store (e.g., Atkinson & Shiffrin, 1971; Glanzer, 1972).

It is therefore perhaps surprising that most current theories of ISR do not provide a detailed account of IFR. For example, currently influential accounts of ISR include the phonological loop model of working memory (Baddeley, 1986, 2000; Baddeley & Hitch, 1974), the Burgess and Hitch (1992, 1999, 2006) model, the primacy model of Page and Norris (1998), the start–end model of Henson (1998), the feature model (Nairne, 1988, 1990), OSCAR (Brown, Preece, & Hulme, 2000), SOB (Farrell & Lewandowsky, 2002; Lewandowsky & Farrell, 2008), and the accounts by Botvinick and Plaut (2006) and by Oberauer and Lewandowsky (2008). None of these current accounts of ISR provide a detailed account of IFR.

Moreover, most current theories of the IFR task do not provide a detailed account of ISR. Influential accounts by Davelaar, Goshen-Gottstein, Ashkenazi, Haarmann, and Usher (2005); the temporal context model (Howard & Kahana, 2002a) and its recent variants (e.g., Polyn, Norman, & Kahana, 2009; Sederberg, Howard, & Kahana, 2008); the account by Laming (2006, 2008, 2009, 2010); the SAM model (Raaijmakers & Shiffrin, 1981); and indeed our own account of IFR (Tan & Ward, 2000) do not explain ISR.

There are a few exceptions, such as the SIMPLE model of Brown, Neath, and Chater (2007); the ACT–R model of Anderson, Bothell, Lebiere, and Matessa (1998); and the LIST PARSE model of Grossberg and Pearson (2008). However, it should be noted that

Geoff Ward and Rachel Grenfell-Essam, Department of Psychology, University of Essex, Colchester, England; Lydia Tan, Department of Psychology, City University, London, England.

Rachel Grenfell-Essam was supported by a 1+3 competitive postgraduate studentship granted by the Economic and Social Research Council, United Kingdom, under the supervision of Geoff Ward. We would like to thank Lydia Allen, Chung Ting Tse, and Jennifer Townsend for help collecting data in Experiments 1 and 2 of this study.

Correspondence concerning this article should be addressed to Geoff Ward, Department of Psychology, University of Essex, Wivenhoe Park, Colchester, England, CO4 3SQ. E-mail: gdward@essex.ac.uk

some of these models differ in the type and number of mechanisms required to underpin the two tasks.

We propose three reasons for the current lack of theoretical integration between these two tasks in the literature. First, there are capacity difficulties in entertaining a limited-capacity STS account of both ISR and the recency effect in IFR. For example, Baddeley and Hitch (1974, 1977) asked participants to study lists of 16 words for free recall whilst concurrently maintaining lists of six digits for ISR. They found that the magnitude of the recency effect in free recall was unaffected by a concurrent six-digit load for ISR, suggesting that active maintenance of six digits for ISR could not be competing for the same limited-capacity STS as the maintenance of the last few words in the list for free recall (a finding since replicated and extended by Bhatarah, Ward, & Tan, 2006). Therefore, theorists are likely to promote different theoretical approaches to recency and ISR if they wish to appeal to STS as an explanation of one of these findings. For example, Baddeley and Hitch concluded that ISR was underpinned by STS but that "working memory, which in other respects can be regarded as a modified STS, does not provide the basis for recency" (Baddeley & Hitch, 1974, p. 81).

Second, early reviews of the IFR and ISR literatures found that ISR was affected by variables such as phonological similarity, word length, presentation rate, and memory load (suggesting a speech-based STS) but that the recency effect in IFR was not particularly sensitive to these variables (see e.g., Baddeley, 1976, p. 182). This differential effect of different variables on ISR and recency further suggested a division between explanations of the recency effect in IFR (not underpinned by STS) and ISR (underpinned by STS).

A third and final reason for the current lack of theoretical integration between theories of IFR and ISR is that the two tasks give rise to highly dissimilar serial position curves. In IFR, explanations of the serial position curve are dominated by explanations of recency effects (e.g., Davelaar et al., 2005; Howard & Kahana, 2002a; Tan & Ward, 2000), whereas in ISR, explanations of the serial position curve are dominated by explanations of primacy effects, the enhanced recall of early list items (e.g., Lewandowsky & Farrell, 2008; Page & Norris, 1998). These contrasting serial position curves offer a difficult challenge to accounts of memory that try to model both tasks, and this difficulty is made all the more transparent in computational models where the calculations and predictions of the models are most explicit.

In short, it is fair to say that it is currently controversial to even consider a common theoretical framework for IFR and ISR (for a recent exchange of views on this issue, see e.g., Brown, Chater, & Neath, 2008; Murdock, 2008).

However, we believe that most previous comparisons between the IFR and ISR tasks have been hindered by the use of different list lengths on the two tasks. Typically, short lists of 5–6 words are used for ISR, whereas much longer lists of 10–40 items are used for IFR. Therefore, apparent differences between the two tasks may have been reflecting differences in the typical list lengths that were used.

Consistent with this possibility, many similarities have recently been found between ISR and IFR when the two tasks have been compared under identical presentation conditions and identical list lengths of 5 or 8 words (Bhatarah, Ward, Smith, & Hayes, 2009; Bhatarah, Ward, & Tan, 2008). For example, (a) words on both tasks are rehearsed and encoded in the same way with little or no effect of test expectancy, even though the output order of the two tasks can be very different; (b) the degree of forward-ordered recall can be similar on the two tasks, even though IFR does not require forward-ordered recall; (c) both tasks are affected by word length and articulatory suppression, variables traditionally associated with ISR; and (d) both tasks are affected similarly by presentation rate, a variable traditionally associated with free recall. These recent findings suggest that (in contrast to many contemporary accounts) common memory mechanisms may operate at encoding in IFR and ISR and that differences in serial position curves could reflect differences in retrieval which can be affected by factors such as list length.

In this research, we seek to examine further the case for greater theoretical integration between ISR and IFR. We again examine both tasks under identical methodological conditions, but in these experiments we additionally manipulate list length from very short lists well within what is typically used in ISR through to longer lists typically reserved for free recall. Central to our analyses is an examination of how increasing the list length affects the output order and the serial position curves in different immediate memory tasks.

General Overview of Methods and Scoring Conventions

Before introducing the individual experiments it may be helpful to provide an overview of the different methods and scoring conventions that we used. Table 1 provides a summary of the different tasks that were used in our four experiments.

Immediate Free Recall (IFR)

Experiments 1 and 2 used IFR. In this task, participants were presented with lists of words, one at a time, at a fairly fast rate of 1 word per second. The number of words in the list (the list length) varied from trial to trial (from between 1 and 15 words) and participants were not told in advance how many words were to be presented on any list. After the last word in the list had been presented, the end of the list was signaled by an auditory cue (Experiment 1), or by an auditory cue and an empty grid on the computer screen, which had as many rows as there had been list items (Experiment 2). This was the participants' cue to start to write down as many words from the list as they could remember on their lined response grids. In IFR, the participants were free to output the words in any order, and they wrote the first word that came to mind in the top row of their response grid and then continued by writing subsequent recalls in successive rows underneath. The scoring method used was free recall scoring (*FR scoring*) in which a word is considered correct (and assigned a 1) if it was recalled at any point during the recall of the list (and assigned a 0 if it was not recalled).

ISR-Free

Experiments 2 and 3 used a task referred to here as *ISR-free* (see also Crowder, 1969; Tan & Ward, 2007; Waugh, 1960). In this task, participants were again presented with sequences of between 1 and 15 words, as before, and at test, they too were required to

Table 1
Overview of the Different Methods Used in the Four Experiments

Experiment and task	*N*	List lengths	Trials	Presentation	Method of response
	Methodological details				
Experiment 1					
IFR	55	1–15	45	Read silently	Written silently
Experiment 2					
IFR	20	1–8, 10, 12, & 15	66	Read aloud	Written whilst spoken aloud
ISR-free	20	1–8, 10, 12, & 15	66	Read aloud	Written whilst spoken aloud
Experiment 3					
ISR-free	20	1–8, 10, 12, & 15	66	Read aloud	Written whilst spoken aloud
Standard ISR	20	1–8, 10, 12, & 15	66	Read aloud	Written silently
Experiment 4					
Standard RoO	15	1–8, 10, 12, & 15	66	Read aloud	Computer mouse silently
Free RoO	15	1–8, 10, 12, & 15	66	Read aloud	Computer mouse silently

Note. IFR = immediate free recall; ISR = immediate serial recall; RoO = reconstruction of order.

write down as many words from the list as they could remember in their lined response grids. However, participants performing ISR-free were asked to write their recalled words in the rows of their response grids that corresponded to the serial positions of the items in the list. Thus, participants were required to write the first list item in the top row of the response grid, the second list item in the second row of the response grid, and so on, such that the last word in the list should be written in the row of the response grid corresponding to the last serial position (the participants could look at the empty response grid on the computer screen to determine the list length of that trial). Participants performing ISR-free were free to write down the words in their respective grid positions in any temporal order. Thus, participants were free if they so wished to start recall by writing the last few words in the list in the last few grid positions, before returning to the top of the grid to continue recall with the early list items. Any temporal order was permitted so long as the participants tried to place the recalled items in the correct grid locations. The scoring system used was *ISR scoring*, in which a word was considered correct (and assigned a 1) if and only if it was placed in the row of the response grid that corresponded to its input position (and assigned a 0 if it was omitted or was not placed in its correct row). Note that the same responses could also be scored using FR scoring in order to determine which words were recalled (irrespective of whether the participant assigned them to their correct serial position).

Standard ISR

Experiment 3 used standard ISR. The presentation of the lists was identical to the two other tasks. The recall instructions were similar to those in ISR-free in that the participants were required to write the first word in the list in the top row of the response grid, the second word in the list in the second row of the response grid, and so on, such that the last word in the list should be written in the row of the response grid corresponding to the last serial position (the participants could again look at the empty response grid on the computer screen to determine the list length of that trial). However, participants performing standard ISR were required to write down their words in the same temporal order as that in which the words were presented. Thus, participants were instructed to start recall with the first word in the list if they could remember it, and then proceed down the grid, attempting to recall the second word, third word, and so forth, in forward serial order. If the participants could not recall one or more words, they were required to continue with the earliest word that they could remember and write it in the row that corresponded to that item. Participants could not return to fill in earlier responses following later responses. The method of scoring was the ISR scoring method, but the responses could also be scored using FR scoring to determine which words were recalled (irrespective of whether the participant correctly assigned them to their serial position).

Standard Reconstruction of Order (Standard RoO)

Experiment 4 used the standard reconstruction of order (RoO) task (see also Healy, 1982; Healy, Fendrich, Cunningham, & Till, 1987; Lewandowsky, Nimmo, & Brown, 2008). The presentation of the words in this task was identical to that used in the other three tasks. In addition, after the last word in the list had been presented, an empty grid appeared on the computer screen containing as many rows as there had been list items. However, unlike the other recall tasks, all the words from the list were also re-presented in a new random order on the computer screen. In the standard RoO task, the participants were required to select the word that they thought was the first word on the list (by clicking on the word using the computer mouse) and then place the selected word in the first location (by clicking at the location of the top row of the response grid on the screen). Once a word was placed in a location, that word appeared in the clicked row in the response grid and was then unavailable to be reselected for subsequent responses. In the standard RoO task, participants were required to reconstruct the list within the computer response grid in a forward serial order, and they could place words only in the highest empty location currently available in the response grid. The only method of scoring is the ISR scoring method, as all the items are re-presented at test and all must be assigned to grid positions.

Free Reconstruction of Order (Free RoO)

Experiment 4 also used the free RoO task, which was identical to the standard RoO task with the exception that the participants were free to place the words into their corresponding locations in the response grid in any temporal order that they wished (see also Lewandowsky, Brown, & Thomas, 2009; Lewandowsky et al., 2008; Nairne, 1991, 1992; Nairne, & Neumann, 1993; Neath, 1997). Again, the only method of scoring is the ISR scoring method, as all the items are re-presented at test and all must be assigned to grid positions.

Scoring Conventions for Analyses of Output Orders

In all four experiments, we were also interested in the order in which the words were recalled or selected. The *probability of first recall* (PFR) refers to the proportion of trials in which the first word recalled (or positioned) was of a particular input serial position (for related analyses, see Hogan, 1975; Howard & Kahana, 1999; Laming, 1999). Previous studies using this measure have examined longer lists in IFR and have typically found that participants initiate IFR with one of the recency items.

We were also interested in the *effect of the first recall on the resultant serial position curves*. Specifically, we partitioned the serial position data on the different tasks by the serial position of the first word recalled. When the first word recalled was that presented in Serial Position 1, we were able to compare the resultant serial position curves with the primacy effects typically obtained in ISR (using ISR scoring). When the first word recalled was one of the last few list items, we were able to compare the resultant serial position curves with the recency effects typically obtained in FR (using FR scoring).

The order of subsequent recalls can be examined by calculating the *conditionalized response probabilities* (CRPs) at different lags, which may be plotted to produce *lag-CRP curves* (e.g., Howard & Kahana, 1999; Kahana, Howard, & Polyn, 2008) for each list length. In a lag-CRP curve, the *x*-axis plots the *lag* between successive pairs of words recalled, which is calculated by subtracting the serial position of the first word of each pair of responses from the serial position of the second word of each pair. Smaller lag values therefore represent recall transitions between words from more similar serial positions, whereas larger lag values represent recall transitions between words from less similar serial positions. In addition, positive lag values represent recall transitions proceeding in a forward direction; negative lag values represent recall transitions proceeding in a backward direction. Of critical interest is the frequency of recall transitions with *lag + 1*, the lag at which the output order of successively recalled pairs is the same as the input order. The *y*-axis plots the CRP, which is calculated by taking the number of transitions actually made by a participant of a given lag during output and dividing this total by the number of opportunities that that participant might reasonably be expected to have had to make such a lag transition. The CRP values control for the reduced opportunities to make transitions at extreme lags (and the increased opportunities to make transitions at small lags) and also assume that it is unreasonable for a participant to recall an item that has already been recalled. Lag-CRP curves have been examined predominantly at longer list lengths in free recall and typically show what is referred to as the *asymmetric lag recency effect* (e.g., Howard & Kahana, 1999; Kahana, 1996), such that there is a preference for transitions to be nearer neighbors than remote neighbors and a greater tendency for transitions to proceed in a forward rather than a backward order. Recently, lag analyses have been performed on free and serial recall of shorter (Bhatarah et al., 2008) and longer lists (Klein, Addis, & Kahana, 2005), where similar degrees of forward-ordered recall have been observed across the two tasks.

There has been some debate in the last few years regarding the additional observation that participants show enhanced CRPs to the extreme positive and extreme negative lags relative to their neighbors (Farrell & Lewandowsky, 2008). Although this was originally considered problematic for the temporal context model of Howard and Kahana (2002a), a subsequent article by Howard, Sederberg, and Kahana (2009) has demonstrated that enhanced CRP values at extreme lags are to be expected if factors that affect the serial position curve are allowed to persist as a retrieval cue throughout recall.

Laming (2006, 2008, 2010) has also claimed that temporal contiguity effects may be limited to pairs of successive items in forward order (lag + 1). According to this view, the asymmetric lag recency effects may arise when these strict forward-ordered contiguity effects are applied to the whole prior sequence of stimuli and rehearsals, which are thought to be preserved in episodic memory in historic order. In support of this claim, when the patterns of rehearsal are observable, there is clear evidence that successive recalls can reflect items that were rehearsed successively (e.g., Laming, 2006; Ward, Woodward, Stevens, & Stinson, 2003) in addition to those that were presented successively.

The lag-CRP plots do not provide a complete analysis of all the sequential dependencies that may occur at recall. It is well known that participants can subjectively organize lists of words in free recall (e.g., Mandler & Dean, 1969; Mandler, Worden, & Graesser, 1974; Pellegrino & Battig, 1974; Tulving, 1962; Wallace, 1970), even for lists of “unrelated” items. The effects of this organization can be seen in patterns of rehearsal (Rundus, 1971) and clustering at recall (Bousfield, 1953; Howard & Kahana, 2002b) and are particularly apparent with repeated presentations of lists (Mandler et al., 1974; Tulving, 1962).

However, the organization on the first presentation of a list is greatly affected by the input order of the list items (Mandler & Dean, 1969; Pellegrino & Battig, 1974; Wallace, 1970), such that lag-CRP plots may provide a good (if not complete) measure of the sequential dependencies that occur for once-presented lists. Moreover, the CRP lag + 1 value can be used to compare directly the degree of forward-ordered recall observed in IFR to that observed in ISR. Nevertheless, it is likely that there are additional interesting sequential dependencies that can be extracted from our data, which can be downloaded from the first author’s website.

Experiment 1

In Experiment 1, we examined only IFR. On each trial, participants were presented with lists of between 1 and 15 words but did not know the length of the list in advance of the cue to recall. The lists were presented visually and read silently at a reasonably fast rate of 1 word per second. We were interested in whether the length of the list affected the order of the words recalled. Specifically, we were interested in knowing whether list length affected the words with which participants would initiate their recall and the extent to which these initial recalls affected the resultant serial position curves.

It is well known that participants tend to start recall with one of the last few list items for IFR of long lists (e.g., Hogan, 1975; Howard & Kahana, 1999; Laming, 1999), but there is at least some evidence that with short lists participants might start recall with the first list item (e.g., Corballis, 1967; Neath & Crowder, 1996).[1] Unfortunately, this observation was only anecdotal in the case of Neath and Crowder (1996) and was limited to 10 participants recalling on only three out of 103 trials in the case of Corballis (1967). Moreover, both of these studies used a single list length of 5 words (Neath & Crowder, 1996) or 5 digits (Corballis, 1967).

Our hypothesis was that if the dissimilarities between IFR and ISR were due at least in part to differences in list length, then the output order and serial position curves of IFR might more closely resemble those of ISR at shorter list lengths. In summary, to the extent that participants started their IFR with the first list item and exhibited extended primacy effects with shorter lists, so we would find evidence for increased similarities between IFR and ISR.

Method

Participants. Fifty-five participants from the University of Essex and City University took part in this experiment.

Materials and apparatus. The materials consisted of a set of 480 words taken from the Toronto Word Pool (Friendly, Franklin, Hoffman, & Rubin, 1982). Subsets of 360 words were randomly selected to be the materials for each individual. The materials were presented in 52-pt Times New Roman font in the center of a computer monitor.

Design. The experiment used a within-subjects design. There were two within-subjects independent variables: list length, with 15 levels (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, and 15), and serial position, with up to 15 levels. The dependent variable was the proportion of words recalled (in any order).

Procedure. Participants were tested individually and informed that they would be shown one practice list of seven words followed by 45 experimental lists of words. The experimental trials were arranged into three blocks of 15 trials. In each block, participants received one trial of each of the 15 different list lengths, but the order of the list lengths within each block was randomized. Each trial started with a warning tone, followed after 3 s by a sequence of between 1 and 15 words presented one at a time in the center of the screen. The presentation rate was 1 word every second, with each word displayed for 0.75 s with an additional 0.25-s interstimulus interval during which the stimulus field was blank. Participants were instructed to read each word silently as it was presented. At the end of the list there was an auditory cue and participants wrote down as many words as they could remember in any order that they wished.

Results

Analyses of serial position curves of all data. We examined the serial position curves using all the data. Figure 1 shows the proportion of words recalled at each of the 15 different list lengths. Consistent with list length effects in IFR, the mean proportion of words recalled decreased monotonically, from .99 (List Length 1) to .30 (List Length 15), $F(14, 756) = 419.3$, $MSE = .008$, $p < .001$. Furthermore, the mean number of words recalled increased from 0.99 (List Length 1) to 4.49 (List Length 15), $F(14, 756) = 124.4$, $MSE = .461$, $p < .001$.

At short list lengths, such as List Length 4, the proportion of words recalled was at a consistently high level ($M = .81$), with no significant effect of serial position, $F(3, 162) = 0.802$, $MSE = .045$, $p > .05$. At longer lists, such as at List Length 8, there was a significant effect of serial position, $F(7, 378) = 10.21$, $MSE = .090$, $p < .001$, reflecting significant recency and primacy effects of approximately equivalent magnitude (all ps $< .01$). At still longer lists, such as at List Length 15, there was also a significant effect of serial position, $F(14, 756) = 20.48$, $MSE = .070$, $p < .001$, reflecting significant recency and primacy effects, but there was significantly greater recall of the recency items than the primacy items (all ps $< .01$).

The probability of first recall (PFR) data. We examined which words from a list were the first to be recalled on each trial. Table 2 shows the PFRs for words at different serial positions in the lists. At short list lengths, participants tended to initiate their recall with the first words in the list (those presented at Serial Position 1), but as the list length increased, so there was an increased tendency to initiate recall with one of the last four words in the list.

Figure 2 clearly illustrates participants' tendency to initiate IFR with the first list item for shorter lists. In addition, as the list length increases, so there is an increasing tendency to initiate IFR with one of the last four list items. There were relatively few trials in which participants initiated recall with one of the other serial positions ("other"), and a handful of trials in which participants either could recall no words from the trial or initiated recall with an error. The proportion of trials in which recall started with Serial Position 1 decreased with increasing list length, $F(14, 756) = 129.9$, $MSE = .048$, $p < .001$. For each participant, we calculated the slope of the function relating how the proportion of trials in which recall started with Serial Position 1 decreased with increasing list length, across sets of List Lengths 1–3, 4–6, 7–9, 10–12, and 13–15. The mean slopes across these five sets (−0.06, −0.17, −0.07, −0.03, and −0.01, respectively) differed significantly from each other, $F(4, 216) = 10.40$, $MSE = .021$, $p < .001$, and confirmed that the function declines most steeply across List Lengths 4–6, where the addition of each extra list item reduced the proportion of trials starting with Serial Position 1 by on average 17%.

We examined whether the tendency to initiate recall with an early list item (observed with Serial Position 1) extended to Serial Positions 2, 3, or 4. We compared the mean proportion of trials in which Serial Positions 2, 3, and 4 were the first word recalled at each list length from 5 to 15 (the means for Serial Positions 2, 3, and 4 were .041, .033, and .036, respectively).[2] A two-way within-subjects analysis of variance (ANOVA) revealed that the proportion of trials starting with one of these early serial positions decreased with increasing list length, $F(10, 540) = 8.36$, $MSE = .010$, $p < .001$, but there was no significant main effect of serial position, $F(2, 108) = 0.71$, $MSE = .017$, $p > .05$, and no significant interaction, $F(20, 1080) = 0.793$, $MSE = .012$, $p > .05$. Thus, although there was a general tendency to initiate recall with

[1] We wish to thank Mike Page for his persistence in pointing out that participants tend to initiate the recall of short lists with the first list item, even in the absence of rehearsal.

[2] We wish to thank Marc Howard for his suggestion for this analysis.

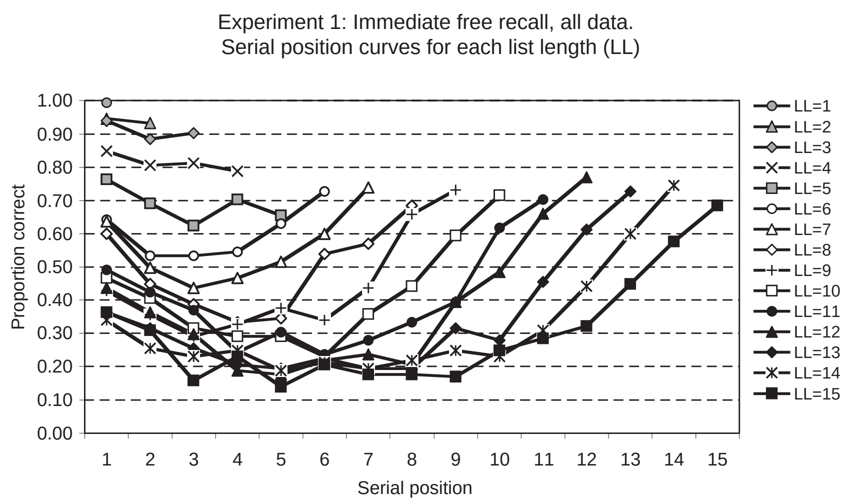


Figure 1. Data from the immediate free recall (IFR) task of Experiment 1 showing the serial position curves: the mean proportion of words recalled from each list length as a function of serial position.

early list items more with shorter lists, there was no evidence for a graded primacy advantage: Participants were equally likely to start their recalls with Serial Positions 2, 3, or 4.

We also examined whether the tendency to start with a late list item (observed with the last serial position, *n*) extended to serial positions $n - 1$, $n - 2$, or $n - 3$. We compared the mean proportion of trials in which serial positions $n - 1$, $n - 2$, or $n - 3$ were the first word recalled at each list length from 5 to 15. A two-way within-subjects ANOVA revealed that the proportion of trials starting with one of these later serial positions increased with list length, $F(10, 540) = 2.45$, $MSE = .024$, $p < .01$; there was a highly significant main effect of serial position, $F(2, 108) = 38.05$, $MSE = .082$, $p < .001$, and a significant interaction, $F(20, 1080) = 1.74$, $MSE = .042$, $p < .05$. Thus, not only was there a general tendency to recall with later items more with longer lists, but there was also evidence for an overall graded recency advantage: Participants were more likely to start their recalls with serial position $n - 1$ than $n - 2$ than $n - 3$. The interaction was because significant recency over serial positions $n - 3$ to $n - 1$ was not universally observed at all list lengths but was observed with List

Table 2
Data From Experiment 1

	List length														
Serial position	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	*164*	*146*	*146*	*116*	*93*	*60*	*45*	*36*	*21*	*23*	*23*	*14*	*21*	*8*	*18*
2		**14**	**11**	**15**	**14**	9	9	7	10	5	5	6	2	3	5
3			**6**	**15**	**14**	**11**	7	7	2	2	8	1	2	4	1
4				**18**	**17**	**7**	**8**	6	9	1	2	4	4	4	3
5					**25**	**35**	**18**	**15**	6	3	4	1	2	1	1
6						**41**	**35**	**24**	**11**	7	4	2	4	1	2
7							**43**	**27**	**21**	**6**	8	2	2	3	0
8								**38**	**34**	**20**	**8**	2	4	5	0
9									**48**	**42**	**18**	**18**	4	8	3
10										**52**	**34**	**20**	**7**	2	5
11											**48**	**44**	**27**	**7**	8
12												**50**	**33**	**20**	**9**
13													**48**	**41**	**25**
14														**55**	**30**
15															**55**
Void	1	4	0	1	0	0	0	0	1	0	0	0	1	1	0
Error on first word	0	1	2	0	2	2	0	5	2	4	3	1	4	2	0
Total	165	165	165	165	165	165	165	165	165	165	165	165	165	165	165

Note. The data show the distribution of the first words recalled on each trial, as a function of the list length and the words' serial position. The italicized values represent the frequency of trials in which the first word recalled was from Serial Position 1, and the bold values represent the frequency of trials in which the first word recalled was from one of the last four serial positions. The values in regular font represent the frequency of trials in which the first word recalled was from one of the other serial positions. Void = no words were recalled on a particular trial.

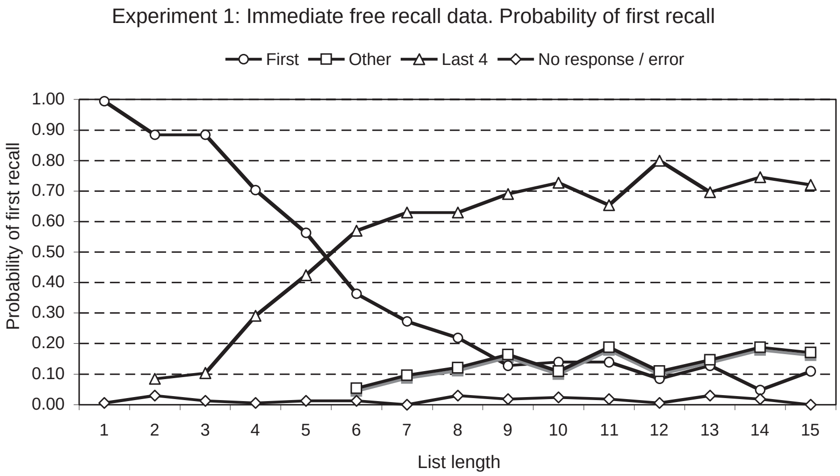


Figure 2. Data from the immediate free recall (IFR) task of Experiment 1 showing the proportion of trials at each list length in which recall initiated with the first word in the list, one of the last four words in the list, or one of the other words in the list. On a small minority of trials, participants either recalled no words from the list (void) or began recall with a word not on the list (error).

Lengths 6, 7, 9, 10, 11, 12, 13, 14, and 15, but not 5 or 8. Thus there was some evidence that for longer list lengths the PFR showed extended recency over a number of serial positions.

The effect of first recall on the serial position curves. We then examined the effect of first recall on the resultant serial position curves. Figure 3A shows the serial position curves for the IFR trials in which the first word recalled was from Serial Position 1 plotted using ISR scoring. Even though the task in Experiment 1 is IFR, it is nonetheless informative to examine recall of trials which start with Serial Position 1 using the ISR scoring method, as this provides an indication of how similar the serial position curves in the IFR task are to those typically obtained in ISR (when the start position and scoring systems are equated). Figure 3B shows the serial position curves for the same data plotted using FR scoring. Finally, Figure 3C shows the serial position curves for the IFR trials in which the first word recalled was from one of the last four serial positions plotted using FR scoring. The shapes of the serial position curves in Figures 3A, 3B, and 3C are strikingly different.

For illustrative purposes, consider performance at List Length 6. Thirty-five participants initiated recall on one or more trials with Serial Position 1 at this list length. Using ISR scoring, the mean recalls across the six serial positions in Figure 3A are 1.00, .55, .38, .26, .18, and .10, and discounting Serial Position 1 (whose mean by definition is 1.00), Serial Positions 2–6 differed significantly, $F(4, 136) = 14.92$, $MSE = .073$, $p < .001$ (pairwise comparisons confirming significant extended primacy effects with ISR scoring).

The corresponding mean recalls across the same trials using FR scoring (see Figure 3B) are 1.00, .67, .59, .59, .46, and .53. Discounting Serial Position 1, Serial Positions 2–6 did not differ significantly, $F(4, 136) = 1.15$, $MSE = .185$, $p > .05$, but the early serial positions remain at a relatively elevated level of recall.

Finally, consider performance by the 43 participants who initiated recall on one or more trials with one of the last four serial positions. Figure 3C shows that the mean recalls across these six serial positions are .48, .32, .44, .54, .79, and .88, means which differed significantly, $F(5, 210) = 16.61$, $MSE = .120$, $p < .001$ (pairwise comparisons confirming one-item primacy, and significant extended recency effects).

Analyses of output order using CRPs. Finally, we examined the extent to which the output orders showed evidence of forward-ordered recall. Figure 4 shows the lag-CRP curves for each list length and provides evidence consistent with the asymmetric lag recency effect (see Howard & Kahana, 1999; Kahana, 1996): There is a preference for transitions to be nearer neighbors than remote neighbors, and there is a far greater tendency for transitions to proceed in a forward order (lag + 1) rather than a backward order (lag – 1), perhaps reflecting the lack of rehearsal assumed in this task (cf. Laming, 2010). A close inspection also suggests that there are also high CRP values at extreme negative lag values, which reflect primacy (Howard et al., 2009).

In order to examine these observations statistically, the lags (and opportunities to output items at different lags) were categorized into eight different lag values: extreme negative lags (the lowest possible lag value at each list length, equivalent to [1 − list length]), remote negative lags (lags [2 − list length] through to –3), –2, –1, +1, +2, remote positive lags (lags + 3 to + [list length − 2]) and extreme positive lags (the highest positive lag value at each list length, equivalent to lag +[list length − 1]). The mean CRPs for the eight respective categories were .21, .06, .13, .19, .51, .16, .09, and .07, which differed significantly, $F(7, 378) = 125.4$, $MSE = .009$, $p < .001$. Pairwise comparisons showed that there was a clear tendency for recall to follow the order at study, as demonstrated by higher CRP values for lag + 1 (.51) than all other lags, including lag – 1 (.19), demonstrating significant asymmetry. There were also significant lag recency effects; the CRP values for lag – 1 (.19) were greater than those for lag – 2 (.13), which in turn were greater than the values for the remote negative lags (.06), and the CRP values for lag + 1 (.51) were greater than those for lag + 2 (.16), which in turn were greater than the values for the remote positive lags (.09). However, the extreme negative

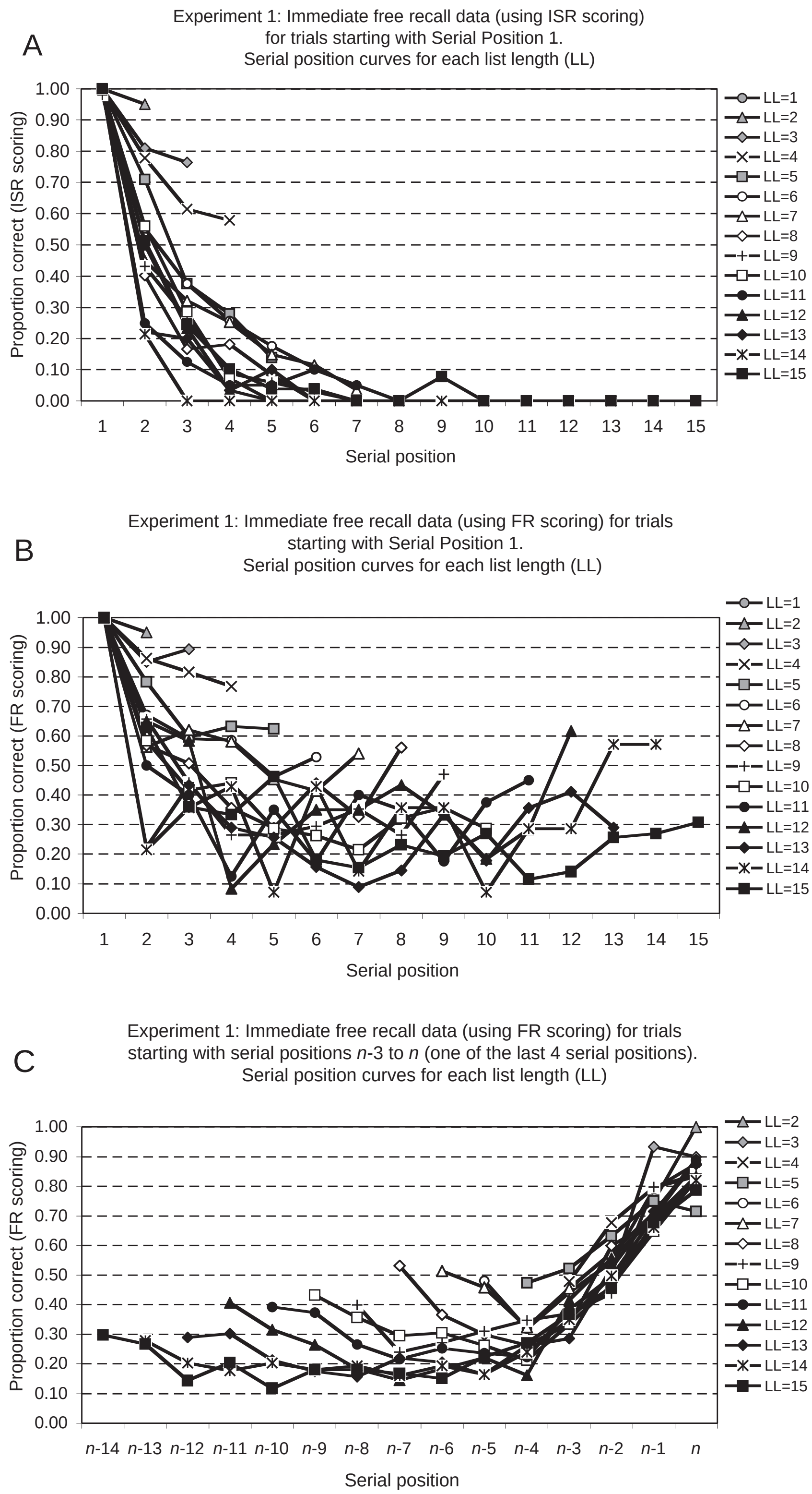


Figure 3. Data from the immediate free recall (IFR) task of Experiment 1 showing the serial position curves for trials that began with Serial Position 1 using (A) immediate serial recall (ISR) scoring and (B) free recall (FR) scoring and (C) for trials that began with one of the last four serial positions using FR scoring. In Panel C, the serial positions have been recency justified.

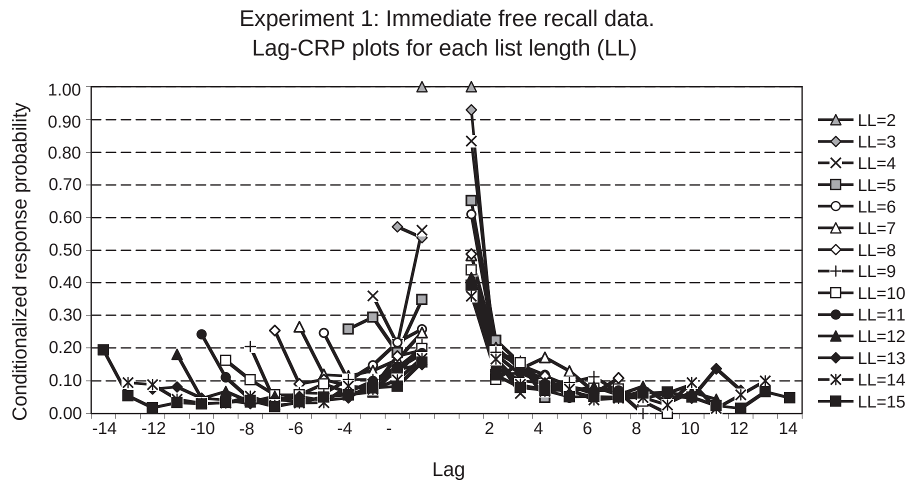


Figure 4. Data from the immediate free recall (IFR) task of Experiment 1 showing lag-CRP (conditionalized response probability) curves for each list length. The lag refers to the difference in serial position between successive words recalled, such that smaller lags reflect the successive recall of words that were presented closer to each other on the list and that positive values reflect pairs of words recalled in the same relative order as at presentation. The CRP represents the mean probability that a word of a particular lag was recalled. It is calculated by dividing the frequency of observed lag transitions by the number of legitimate opportunities in which words at each lag could be recalled.

lag CRP values (.21) were also significantly higher than all but the closest transitions, indicating a high tendency to transition between the last and the first list item.

An inspection of Figure 4 also shows that the CRP values for lag + 1 decreased with increasing list length, a finding shown more clearly in Figure 5. Fifty-one participants had valid lag + 1 CRP values at all 14 list lengths (List Lengths 2–15), and these values differed significantly from each other, $F(13, 637) = 42.53$, $MSE = .051$, $p < .001$. At shorter list lengths, the majority of transitions between successive words output were in the same order as that in which the words were presented (lag + 1), even though the participants were free to recall in any order. The lag + 1 value decreased with increasing list length but did not reduce to near chance levels (approximately the reciprocal of the list length – 1) but stayed at a relatively high value, close to .4.

Discussion

When participants were presented with short list lengths of up to 3 words for IFR, they started their recall with the first word in the list and continued in forward serial order; that is, their IFR resembled ISR. As the list length was increased from lengths of 4–9

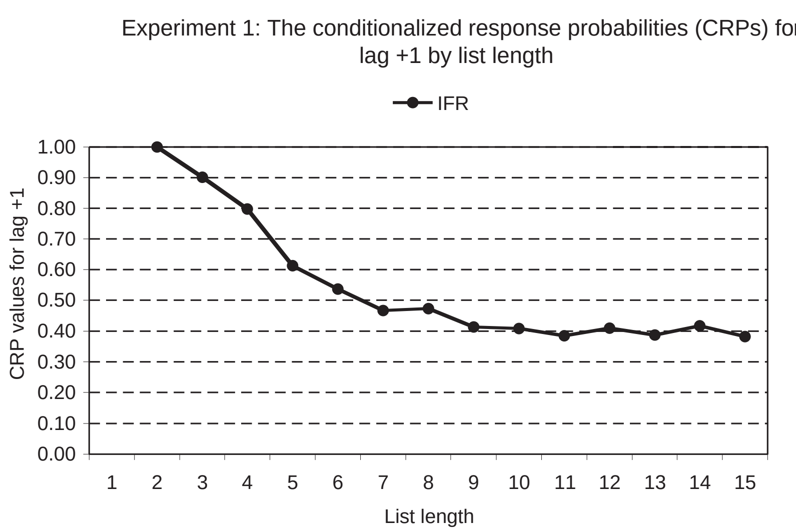


Figure 5. Data from Experiment 1 showing the proportion of lag + 1 responses as a function of list length. CRP = conditionalized response probability; IFR = immediate free recall.

words, so there was an increased tendency to initiate recall with one of the last four words (as is more typical in IFR).

In addition, the first word recalled helped determine the shape of the serial position curve: Heightened early list performance and reduced recency effects were observed when the first word recalled was from Serial Position 1, and more extended recency effects and reduced primacy effects were observed when the first word recalled was from one of the last four serial positions. In all recall, there was a clear tendency to recall in a forward direction, but this tendency was greater for shorter lists.

It appears therefore that at least some of the apparent differences between the output orders and serial position curves observed in IFR and ISR might indeed reflect the typical differences in the list lengths that are used in the two tasks—when IFR was performed with long lists typically used in free recall, participants recalled the last items first and there was enhanced recency and reduced primacy; but when IFR was performed with shorter lists such as those typically used in ISR, then ISR-like findings were observed, with an increased tendency to recall spontaneously from the start of the list, leading to enhanced early list performance on these trials.

Experiment 2

Experiment 1 provided initial evidence supporting the call for greater theoretical integration between IFR and ISR: Some ISR-like features of recall were observed when IFR was performed with shorter, ISR-like list lengths. In Experiment 2, we attempted to replicate these findings and extend them to examine whether ISR might additionally show some free recall-like features of recall when performed under free recall-like list lengths.

An immediate concern with such an endeavor was to consider the best way to examine ISR at such a wide range of list lengths. Our solution in Experiment 2 was to use a variant of the ISR task used by Tan and Ward (2007, see also Crowder, 1969; Waugh, 1960) that we refer to as ISR-free. In the ISR-free task, participants were presented with a list of words one at a time, and at the end of the list, they were required to recall the list items in the correct serial position by writing the list items in the appropriate position in a lined response grid. In Experiment 2, the paper response grid always consisted of two columns of 15 rows; the first column contained the numbers 1 to 15 in increasing order, and the second column was empty. Thus, the participants were required to write down the words from the list in the rows corresponding to each item's serial position, such that the first word in the list should be written in the first row of the grid, the second word should be written in the second row, and so on. Unlike standard ISR, participants in the ISR-free task were free to write their spatially ordered responses in the response grids in any temporal order that they liked. For example, they were free to start their recall by writing the last three words from the list in the corresponding rows before returning to the top of the grid to continue recall with earlier list items.

In Experiment 2, one group of participants performed IFR and a second group performed ISR-free. All were presented with a total of 66 lists of between 1 and 15 words. In order to maximize the number of trials per list length within an experimental session, List Lengths 9, 11, 13, and 14 were excluded, such that participants undertook six trials each of 11 different list lengths (lists of 1, 2, 3, 4, 5, 6, 8, 10, 12, and 15 words). As in Experiment 1, the words were presented visually at a reasonably fast rate of 1 word per second. Unlike Experiment 1, the participants were required to read each word aloud as the words were presented. At test, an empty response grid appeared on the screen that had the exact number of rows (between 1 and 15) appropriate for that trial, so that participants could see the number of items that had been presented and so knew which of the 15 rows on their paper response grid they should complete. In the IFR condition, participants were free to recall the words in any order—they simply wrote down the words from the top to the bottom of their response grids. In the ISR-free condition, the participants wrote the words in the positions in the grid corresponding to each word's serial position, but they were free to write down the words in any temporal order that they liked. In both conditions, the participants spoke out loud the words that they were writing down and their responses were recorded using a tape recorder to code the output order of the items.

As in Experiment 1, we were interested in the order in which participants would initiate their recall at different list lengths and the extent to which these initial recalls affected the resultant serial position curves. Our hypothesis was that if the dissimilarities between IFR and ISR were due primarily to differences in list length, then when the list length in ISR-free was increased to list lengths typically associated with IFR the output orders and serial position curves found in both IFR and ISR-free tasks might both tend to resemble those typically observed with IFR with long lists. To the extent that participants switched from initiating recall with the first word to initiating recall with one of the last list items on both tasks with increasing list length, so we would find evidence for increased similarities between IFR and this variant of ISR.

Method

Participants. Forty volunteers from the University of Essex participated in this experiment. None had participated in Experiment 1.

Materials and apparatus. The materials for each participant consisted of 438 words randomly selected from a set of 480 words taken from the Toronto Word Pool (Friendly et al., 1982). The materials were presented in 52-pt Times New Roman font in the center of a computer monitor. Participants received a response booklet of 66 response grids, each of which contained two columns and 15 rows. The first (narrow) column contained the numbers 1 to 15 in ascending order, the second (wide) column allowed room for participants to write their responses.

Design. The experiment used a mixed design. The type of task was manipulated between subjects. There were two within-subjects independent variables: list length, with 11 levels (1, 2, 3, 4, 5, 6, 7, 8, 10, 12, and 15), and serial position, with up to 15 levels. The dependent variable for IFR was the proportion of words recalled (in any order), and the dependent variable for ISR-free was the proportion of words recalled (in the correct serial position).

Procedure. Participants were presented with a total of 66 trials consisting of two blocks of IFR trials or two blocks of ISR-free trials. In each block there were 33 experimental trials, which consisted of three trials of each of the 11 different list lengths. The order of the list lengths within each block was randomized. Participants were tested individually, and the first block was preceded by task-specific instructions and a practice list of List Length 7 on that task.

Each trial started with a warning tone, followed after 3 s by a sequence of between 1 and 15 words presented one at a time in the center of the screen. The presentation rate was 1 word every second, with each word displayed for 0.75 s with an additional 0.25-s interstimulus interval during which the stimulus field was blank. Participants were instructed to read each word aloud as it was presented. At the end of the list there was an auditory cue and an empty grid was displayed on the computer screen. The computer grid resembled the paper response grid and contained two columns but only as many rows as there had been words on the list. Like the paper response grids, the first (narrow) column contained the numbers 1 to list length in ascending order, and the second (wide) column was left blank.

The participants wrote down as many words as they could remember in their lined response grids whilst saying out loud what they were writing down. Participants spoke their recalls out loud, which were recorded with a tape recorder so that the temporal order of output could be subsequently coded off-line. In IFR, the participants were free to write down the words in any temporal order that they wished and filled their response grids from the top of the grid. In ISR-free, the participants were free to write down the words in any temporal order that they wished, but they were instructed to write down the words in the position in the response grid that corresponded to each word's serial position.

Results

Analyses of serial position curves of all data. Figure 6A shows the proportion of words recalled at each of the 11 different

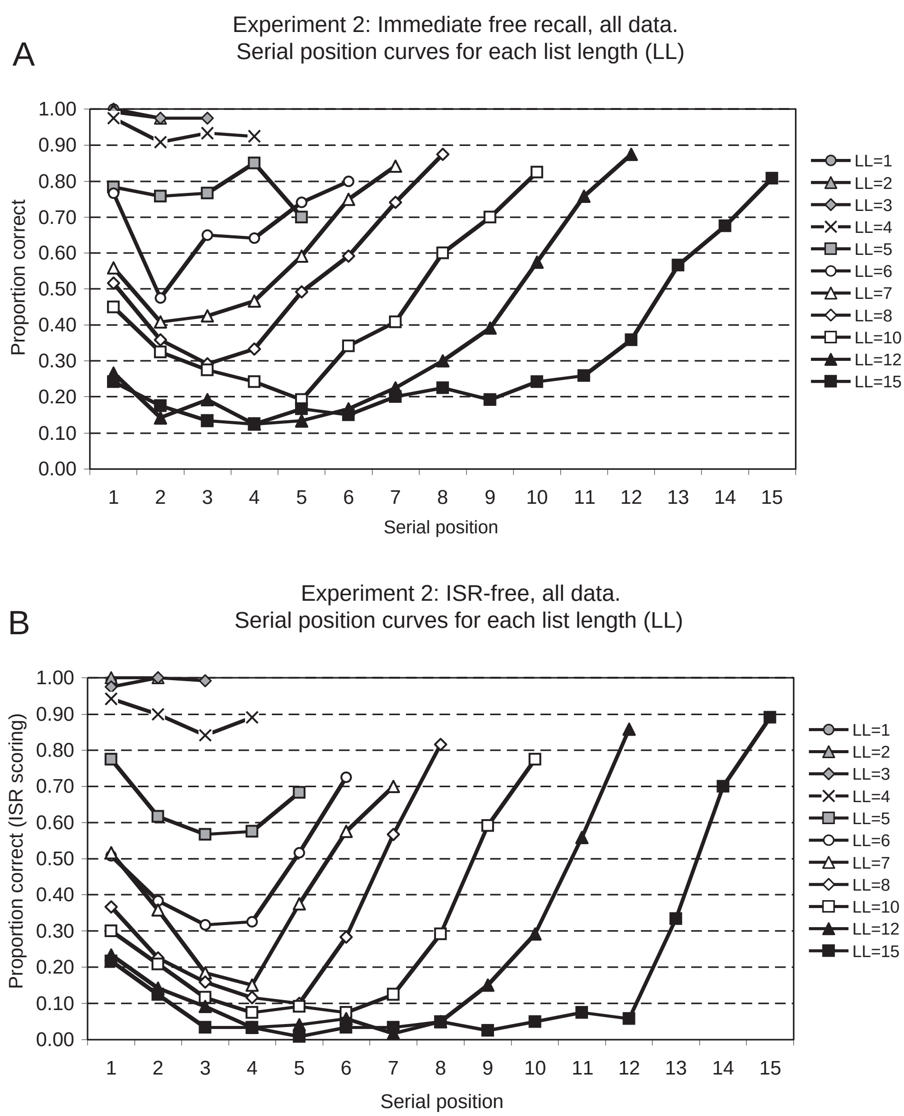


Figure 6. Data from Experiment 2 showing (A) the serial position curves for the free recall task and (B) the ISR-free task for each list length. ISR = immediate serial recall.

list lengths in the IFR task, using FR scoring. Consistent with list length effects in IFR, the mean proportion of words recalled decreased monotonically, from 1.00 (List Length 1) to .30 (List Length 15), $F(10, 190) = 387.2$, $MSE = .004$, $p < .001$, and the mean number of words recalled increased from 1.00 (List Length 1) to 4.52 (List Length 15), $F(10, 190) = 94.25$, $MSE = .263$, $p < .001$.

At short list lengths, such as List Length 4, the proportion of words recalled was consistently at a high level ($M = .94$), with no significant effect of serial position, $F(3, 57) = 2.26$, $MSE = .007$, $p > .05$. At longer lists, such as at List Length 8, there was a significant effect of serial position, $F(7, 133) = 17.27$, $MSE = .049$, $p < .001$, reflecting significant primacy and extended recency effects, but there was significantly greater recall of the recency items than the primacy items.

Figure 6B shows the proportion of words recalled in the ISR-free task at each of the 11 different list lengths in the correct serial position (ISR scoring). The mean proportion of words recalled decreased monotonically from .99 (List Length 1) to .18 (List Length 15), $F(10, 190) = 389.9$, $MSE = .006$, $p < .001$, but the mean number of words recalled increased from 0.99 (List Length 1) to 3.58 (List Length 4) but then declined to 2.67 (List Length 15), $F(10, 190) = 32.53$, $MSE = .278$, $p < .001$.

At short list lengths, such as List Length 4, the proportion of words recalled was at a high level ($M = .89$), but there was nonetheless a significant effect of serial position, $F(3, 57) = 3.51$, $MSE = .010$, $p < .05$, reflecting a significant difference between Serial Positions 1 and 3. At longer list lengths, such as at List Length 8, there was a significant effect of serial position, $F(7, 133) = 24.45$, $MSE = .051$, $p < .001$, reflecting significant primacy and extended recency effects, but there was significantly greater recall of the recency items than the primacy items.

Finally, we compared the proportion of words recalled (using FR scoring) in the IFR task with the proportion of words recalled in the ISR-free task (also using FR scoring). The means for both the IFR and the ISR-free data can be found later in Table 6. A 2 (task) × 11 (list length) mixed ANOVA revealed a nonsignificant main effect of task, $F(1, 38) = 0.014$, $MSE = .004$, $p > .05$, a highly significant main effect of list length, $F(10, 380) = 777.5$, $MSE = .029$, $p < .001$, and a nonsignificant interaction, $F(10, 380) = 1.19$, $MSE = .029$, $p > .05$. Thus, the mean proportions of words recalled at each list length in the IFR conditions were very similar to the mean proportions of words recalled at each list length in the ISR-free conditions using FR scoring.

PFR data. Table 3 shows the PFR data for lists of different lengths in the IFR and ISR-free tasks. For both tasks, the serial position refers to the input position of the word that was first recalled. At short list lengths, participants tended to initiate their recall with the first words in the list (those presented at Serial Position 1), but as the list length increased, so there was an increased tendency to initiate recall with one of the last four words in the list. A similar pattern is apparent for both tasks.

Figure 7 shows a summary of the PFR data for the IFR task (see Figure 7A) and for the ISR-free task (see Figure 7B). We examined the proportion of trials in which Serial Position 1 was output first for the two tasks. This revealed striking similarities between the two tasks: There was a nonsignificant main effect of task, $F(1, 38) = 0.695$, $MSE = .201$, $p > .05$, a highly significant main effect of list length, $F(10, 380) = 200.1$, $MSE = .029$, $p < .001$, and a nonsignificant interaction, $F(10, 380) = 1.23$, $MSE = .029$, $p > .05$.

For the IFR data, the mean slopes across List Lengths 1–3, 4–7, and 8–15 were −0.02, −0.18, and −0.04, respectively. The corresponding slopes for the ISR-free data were −0.01, −0.18, and −0.02. These slopes were highly similar: There was a nonsignificant main effect of task, $F(1, 38) = 0.550$, $MSE = .003$, $p > .05$, a highly significant main effect of list length, $F(2, 76) = 73.19$, $MSE = .005$, $p < .001$, and a nonsignificant interaction, $F(2, 76) = 0.24$, $MSE = .005$, $p > .05$. Thus, for both tasks, the addition of each extra list item between List Lengths 4–7 decreased the proportion of trials starting with Serial Position 1 by, on average, 18%.

We examined whether the tendency to initiate recall with an early list item (observed with Serial Position 1) extended to Serial Positions 2, 3, and 4. We compared the mean proportion of trials in which Serial Positions 2, 3, and 4 were the first word recalled at each list length from 5 to 15 on the two tasks. For the IFR data, the mean proportions of words first recalled from Serial Positions 2, 3, and 4 were .023, .044, and .031, respectively. The corresponding data for the ISR-free task were .017, .022, and .031, respectively. A 2 (task) × 7 (list length) × 3 (serial position) ANOVA revealed a nonsignificant main effect of task, $F(1, 38) = 1.51$, $MSE = .013$, $p > .05$, a significant main effect of list length, $F(6, 228) = 11.02$, $MSE = .005$, $p < .001$, and a significant main effect of serial position, $F(2, 76) = 3.39$, $MSE = .004$, $p < .05$. No interactions were significant. The probability of starting recall with Serial Positions 2, 3, and 4 decreased with increasing list length and critically was *less* for Serial Position 2 than Serial Positions 3 and 4. There was therefore no tendency for graded primacy within the PFR for the two tasks.

The effect of first recall on the serial position curves. We then examined the effect of the first recall on the resultant serial position curves. Figure 8A shows the serial position curves for the IFR trials in which the first word recalled was from Serial Position 1, plotted using ISR scoring. Figure 8B shows the serial position curves for those same trials plotted using FR scoring. Finally, Figure 8C shows the serial position curves for the IFR trials in which the first word recalled was from one of the last four serial positions plotted using FR scoring. Replicating the findings of Experiment 1, the shapes of the serial position curves of the IFR data in Figures 8A, 8B, and 8C are strikingly different. The corresponding serial position curves for the ISR-free data are shown in Figures 8D, 8E, and 8F.

We again limit our analyses to performance at a representative list length, List Length 6. We first compared the List Length 6 data on the two tasks (using ISR scoring) from trials in which recall initiated with Serial Position 1 (see Figures 8A and 8D). Nineteen participants in the IFR task (mean ISR scores across the six serial positions were 1.00, .44, .26, .19, .12, and .08) and 17 participants from the ISR-free task (mean ISR scores across the six serial positions were .96, .57, .40, .23, .25, and .55) contributed to this analysis. Ignoring Serial Position 1, a 2 (task) × 5 (serial position: 2–6) mixed ANOVA revealed a significant main effect of task, $F(1, 34) = 9.39$, $MSE = .155$, $p < .01$, a highly significant main effect of serial position, $F(4, 136) = 8.79$, $MSE = .065$, $p < .001$, and a significant interaction, $F(4, 136) = 3.84$, $MSE = .065$, $p < .01$. However, pairwise comparisons revealed that the significant main effect of task and the significant Task × Serial Position

Table 3
Data From Experiment 2 for IFR and ISR-Free

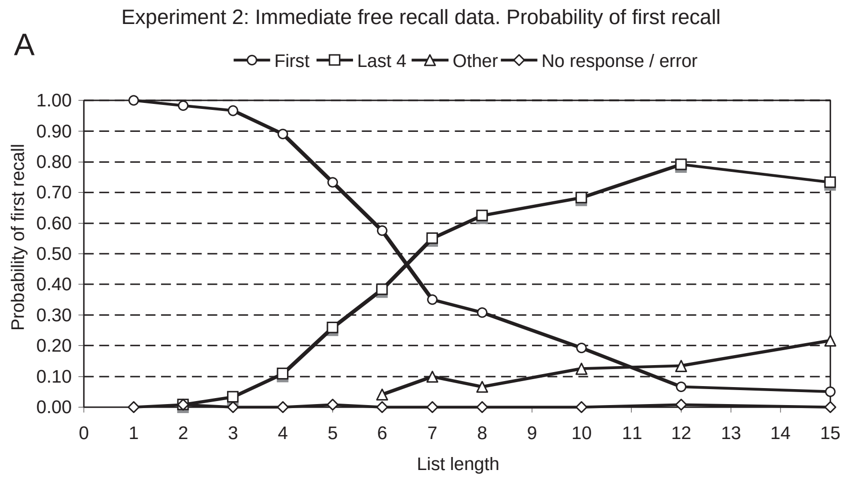
	List length										
Serial position	1	2	3	4	5	6	7	8	10	12	15
						IFR					
1	*120*	*118*	*116*	*107*	*88*	*69*	*42*	*37*	*23*	*8*	*6*
2		**1**	**1**	**6**	**5**	5	6	1	0	1	1
3			**3**	**2**	**10**	**12**	6	4	2	3	0
4				**5**	**7**	**3**	**4**	3	6	1	2
5					**9**	**13**	**13**	**13**	3	0	0
6						**18**	**26**	**11**	4	3	3
7							**23**	**16**	**11**	3	3
8								**35**	**14**	5	5
9									**22**	**7**	3
10									**35**	**11**	4
11										**32**	5
12										**45**	**7**
13											**17**
14											**14**
15											**50**
Void	0	0	0	0	0	0	0	0	0	0	0
Error on first word	0	1	0	0	1	0	0	0	0	1	0
Total	120	120	120	120	120	120	120	120	120	120	120
						ISR-free					
1	*119*	*120*	*116*	*102*	*81*	*50*	*40*	*22*	*18*	*12*	*7*
2		**0**	**3**	**4**	**6**	3	4	0	0	1	0
3			**1**	**2**	**4**	**8**	4	2	0	0	0
4				**12**	**6**	**9**	**5**	5	1	0	0
5					**23**	**14**	**15**	**4**	2	1	0
6						**36**	**15**	**15**	3	1	0
7							**37**	**21**	**8**	0	0
8								**51**	**13**	2	0
9									**22**	**12**	1
10									**53**	**9**	2
11										**22**	2
12										**59**	**5**
13											**19**
14											**18**
15											**66**
Void	0	0	0	0	0	0	0	0	0	1	0
Error on first word	1	0	0	0	0	0	0	0	0	0	0
Total	120	120	120	120	120	120	120	120	120	120	120

Note. The data show the distribution of the first words recalled on each trial, as a function of the list length and the words' serial position. The italicized values represent the frequency of trials in which the first word recalled was from Serial Position 1, and the bold values represent the frequency of trials in which the first word recalled was from one of the last four serial positions. The values in regular font represent the frequency of trials in which the first word recalled was from one of the other serial positions. IFR = immediate free recall; ISR = immediate serial recall; void = no words were recalled on a particular trial.

interaction were entirely due to the enhanced one-item recency observed in the ISR-free task relative to the IFR task. When the recall of the Serial Positions 2–5 were compared using a 2 × 4 mixed ANOVA, there were no significant differences between the two tasks, nor was there a significant interaction. Thus, the serial position curves for List Length 6 were fairly similar for the two tasks across Serial Positions 1–5 using ISR scoring when recall started with Serial Position 1. There were extended primacy effects in both tasks, but there was additionally a one-item recency effect at Serial Position 6 in the ISR-free data.

We then compared the same List Length 6 data but analyzed recall using FR scoring (see Figures 8B and 8E). The mean recalls for the IFR task (see Figure 8B) were 1.00, .59, .61, .55, .63, and .72, whereas for the ISR-free task (see Figure 8F) the mean recalls were 1.00, .61, .57, .43, .52, and .67. A 2 (task) × 5 (serial position: 2–6) mixed ANOVA revealed a nonsignificant main effect of task, $F(1, 34) = 1.28$, $MSE = .137$, $p > .05$, a nonsignificant main effect of serial position, $F(4, 136) = 1.81$, $MSE = .108$, $p > .05$, and a nonsignificant interaction, $F(4, 136) = 0.28$, $MSE = .108$, $p > .05$. Thus the serial position curves for List Length 6 were fairly similar for the two tasks using FR scoring and showed consistently elevated recall performance when recall started with Serial Position 1.

Finally, we compared performance by the 17 participants who initiated recall on one or more trials with one of the last four serial positions in the IFR task (see Figure 8C) with the 18 participants who started one or more trials with one of the last four serial positions in the ISR-free task (see Figure 8F). The mean recalls



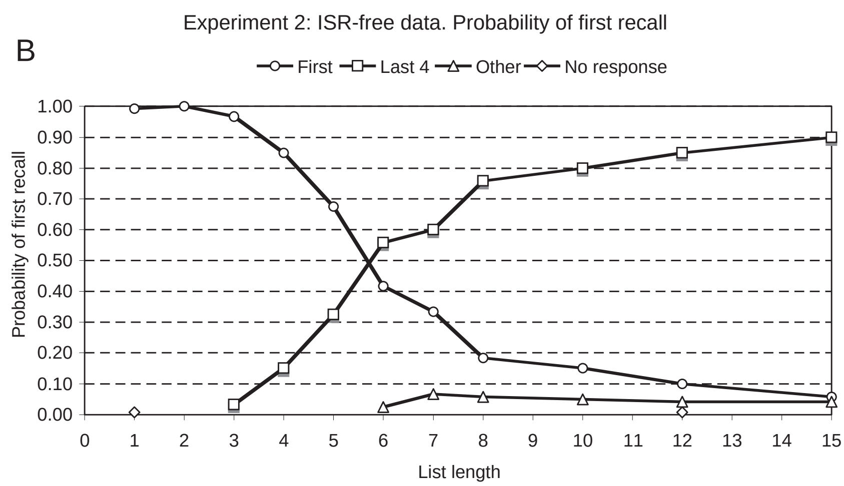


Figure 7. Data from (A) the immediate free recall (IFR) task and (B) the ISR-free task of Experiment 2 showing the proportion of trials at each list length in which recall initiated with the first word in the list, one of the last four words in the list, or one of the other words in the list. On a small minority of trials, participants either recalled no words from the list (void) or began recall with a word not on the list (error). ISR = immediate serial recall.

across the six serial positions for the IFR task in Figure 8C were .48, .17, .67, .71, .84, and .86. The mean recalls across the six serial positions for the ISR-free task in Figure 8F were .32, .31, .55, .68, .87, and .95. A 2 (task) × 6 (serial position) mixed ANOVA revealed a nonsignificant main effect of task, $F(1, 33) = 0.039$, $MSE = .103$, $p > .05$, a significant main effect of serial position, $F(5, 165) = 27.89$, $MSE = .084$, $p < .001$, and a nonsignificant interaction, $F(5, 165) = 1.44$, $MSE = .084$, $p > .05$. Thus the serial position curves for List Length 6 were fairly similar for the two tasks using FR scoring and showed extended recency when recall started with one of the last four positions.

Analyses of output order using CRPs. Finally, we examined the extent to which the subsequent outputs showed evidence of forward-ordered recall. For both tasks, the lag refers to the difference in serial positions based on the input positions of the words that were recalled. Figure 9 plots the CRP values for lag + 1 responses as a function of list length for IFR and ISR-free. Note that neither task requires participants to output successive items in the same order as at input. The CRP values for lag + 1 responses were analyzed by a 2 (task: IFR and ISR-free) × 10 (list length: 2, 3, 4, 5, 6, 7, 8, 10, 12, and 15) mixed ANOVA, which revealed a significant main effect of task, $F(1, 38) = 6.65$, $MSE = .086$, $p < .05$, a significant main effect of list length, $F(9, 342) = 79.88$, $MSE = .024$, $p < .001$, and a nonsignificant interaction, $F(9, 342) = 1.54$, $MSE = .024$, $p > .05$. Thus, the proportion of forward-ordered recall decreased with increasing list length for both tasks, but there was more forward-ordered recall in the ISR-free task than the free recall task. Note that even at the longer list lengths there remains considerable forward-ordered recall (CRP values for lag + 1 remain greater than .40) for both tasks, even though there is no formal requirement to output in forward order in the two tasks.

Discussion

Experiment 2 provided further evidence for the need for greater theoretical integration between IFR and ISR. Replicating the main

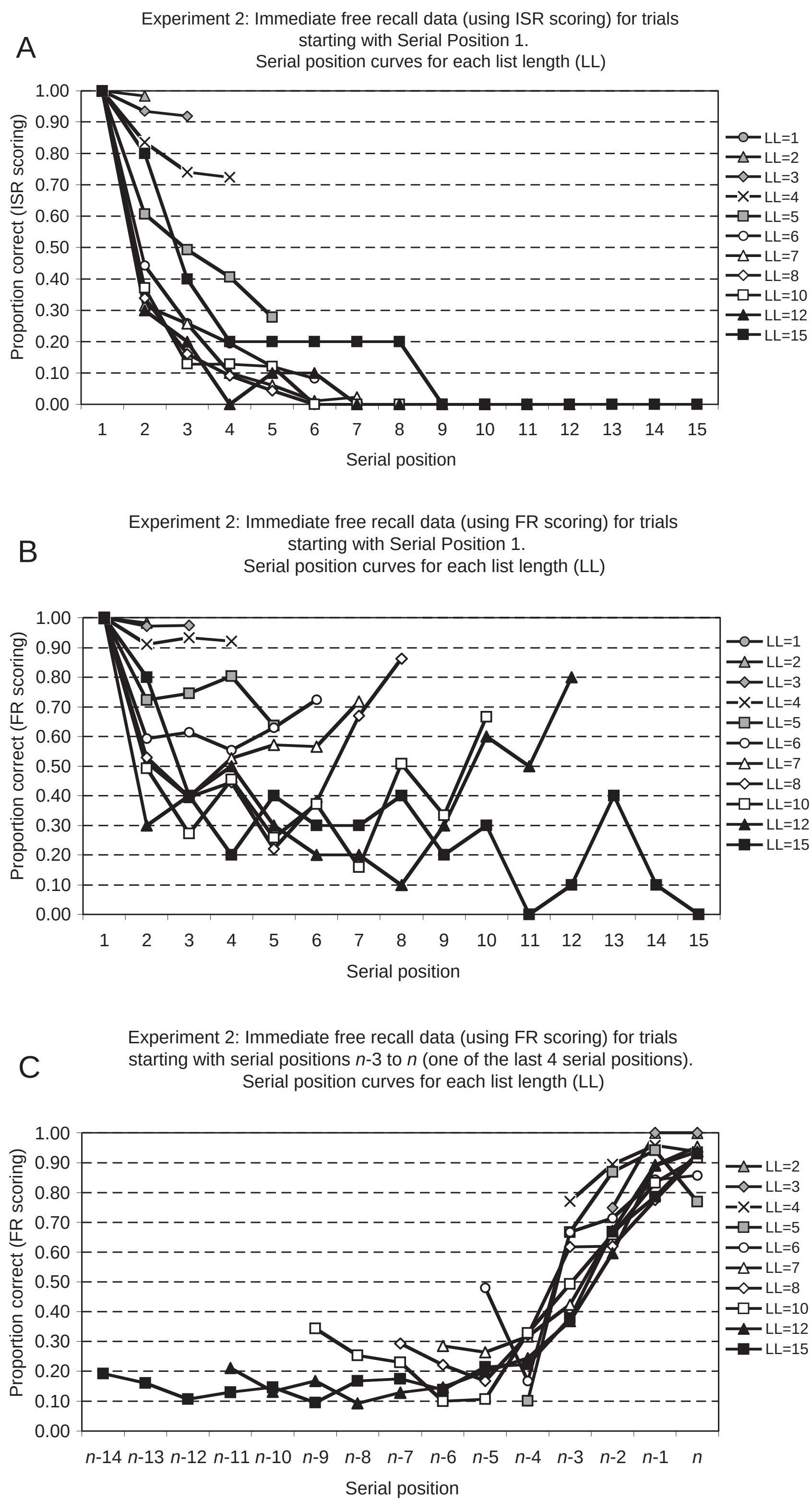


Figure 8. Data from Experiment 2 showing serial position curves for free recall trials that began with Serial Position 1 using (A) immediate serial recall (ISR) scoring and (B) free recall (FR) scoring, and (C) for trials that began with one of the last four serial positions using FR scoring. Panels D–F show serial position curves for the ISR-free data for trials that began with Serial Position 1 using (D) ISR scoring and (E) FR scoring and (F) for trials that began with one of the last four serial positions using ISR scoring. In Panels C and F, the serial positions have been recency justified.

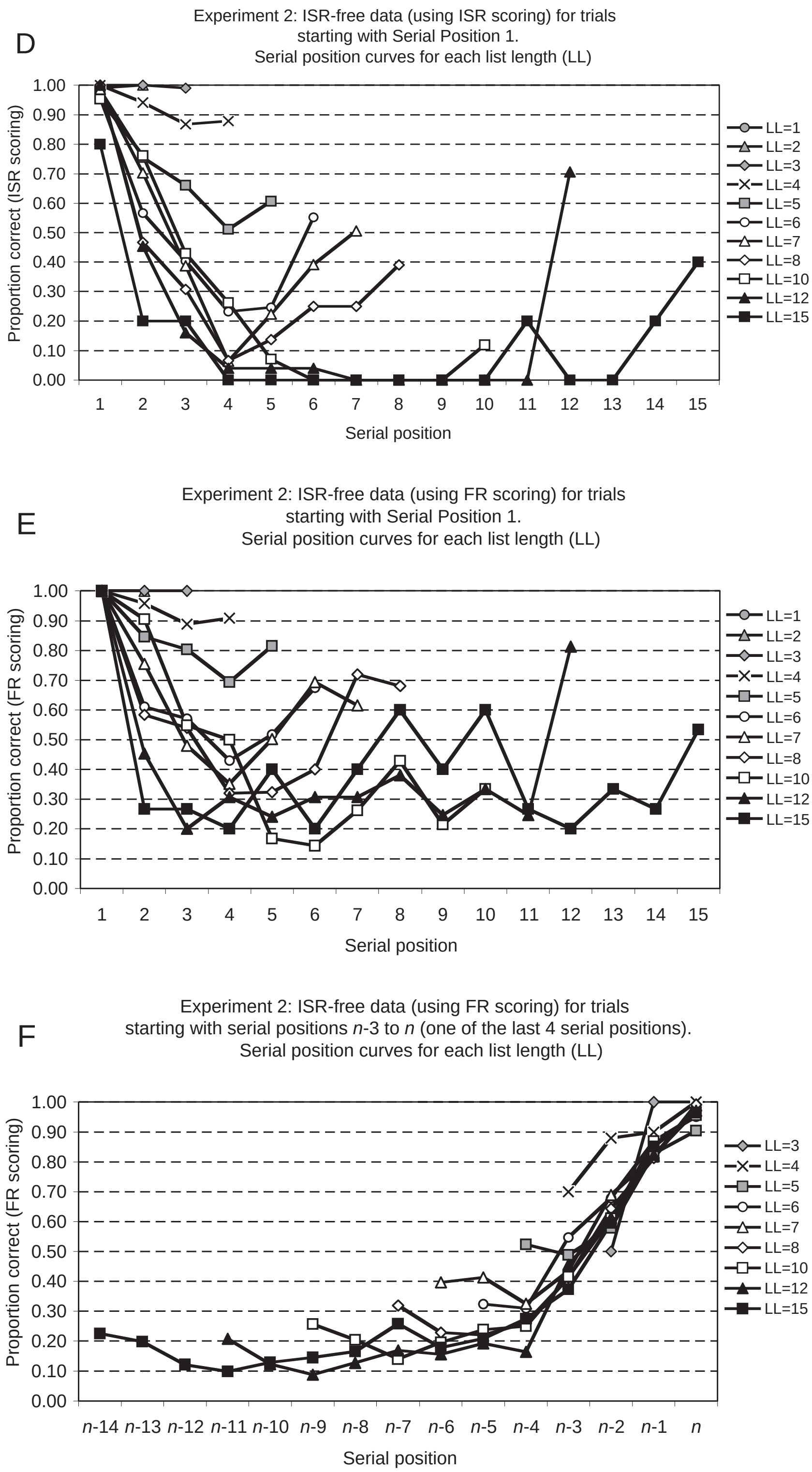


Figure 8. (continued)

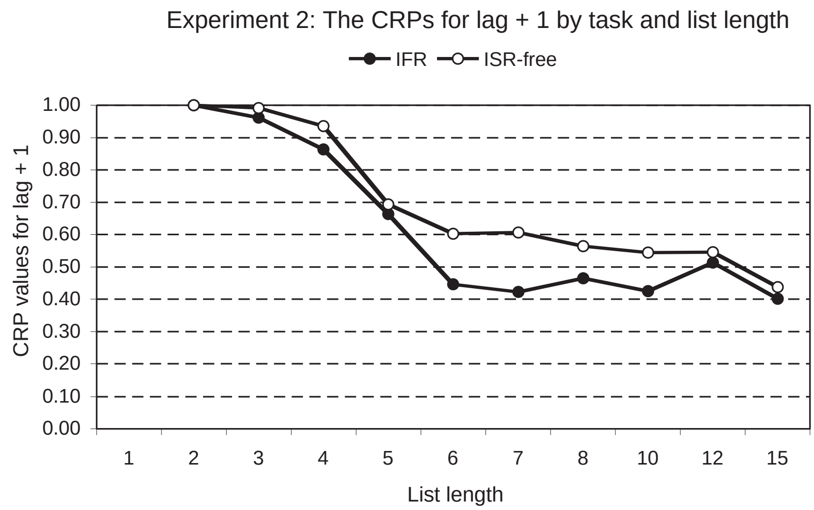


Figure 9. Data from Experiment 2 showing the proportion of lag + 1 responses as a function of list length and task. IFR = immediate free recall; ISR = immediate serial recall; CRP = conditionalized response probability.

findings of Experiment 1, Experiment 2 found ISR-like features in IFR when it was performed with shorter ISR-like list lengths. In addition, Experiment 2 found some free recall-like features in serial recall when ISR-free was performed under free recall-like list lengths.

Specifically, in both tasks, participants started their recalls with the first word on the list when the list was short, but as the list length increased, so there was an increased tendency to start recall with one of the last four words. Moreover, in both tasks, the initial response had a large effect on the subsequent serial position curves: Heightened recall of early list positions occurred with limited recency when recall began with Serial Position 1, but extended recency with limited primacy was observed when recall began with one of the last four items.

The highly similar levels of recall using FR scoring on the two tasks might make one wonder whether the ISR-free task was performed by first performing an IFR phase to generate the words, followed by a second phase in which the words were allocated to their correct serial position. In support of this idea, the two tasks exhibit very similar patterns of overall performance at different list lengths, similar PFR curves, and similar serial position curves for List Length 6 using FR scoring. In addition, for both tasks the lag + 1 values decreased with increasing list length, suggesting similarities in the output orders of the two tasks at shorter list lengths. However, at longer list lengths there were more forward-ordered responses in the ISR-free task than the IFR task, suggesting that the task requirement to position the words in the correct grid position in the ISR-free task encouraged participants to retrieve longer runs of ordered recalls. Thus, when participants recall short lists, ISR-free can be considered to be IFR followed by allocating the items to a serial position, but at longer list lengths, the task requirement to allocate the words into serial positions seems to help guide the generation of the list items in the ISR-free task (or at least influence the recall order).

Experiment 3

Experiment 2 compared IFR performance with performance on ISR-free, a variant of ISR. An advantage of using ISR-free is that we can see which words participants would choose to output first in a serial recall test. It therefore speaks to which words are most accessible at the time of test during ISR (see Tan & Ward, 2007). However, it remains to be seen whether the list length effects observed on the output orders and serial position curves using ISR-free would also be found when the participants were instructed to perform a more standard ISR task.

In Experiment 3, the effects of list length on the output orders and serial position curves were examined using a more standard ISR task as well as the ISR-free task (used by Tan & Ward, 2007, and Experiment 2). A total of 40 participants were presented with 66 lists of words for serial recall. Half the participants received the ISR-free instructions, and the other half received the more standard ISR instructions. Specifically, we adopted the instructions used by Golomb, Peelle, Addis, Kahana, and Wingfield (2008), who examined serial recall of 10-word lists. Golomb et al. encouraged participants to begin recall with the first list item; however, because of the supraspan list length, they instructed participants that if they were unable to retrieve the first item, they should begin their recall with the earliest item that they could remember.

Method

Participants. Forty volunteers from the University of Essex participated in this experiment. None had participated in any of the earlier two experiments.

Materials and apparatus. The materials were the same as those used in Experiment 2.

Design. The experiment used a mixed design. The type of task was manipulated between subjects with two levels: standard ISR and ISR-free. There were also two within-subjects independent variables: list length, with 11 levels (1, 2, 3, 4, 5, 6, 7, 8, 10, 12, and 15), and serial position, with up to 15 levels. The dependent variable was the proportion of words recalled in the correct serial position.

Procedure. The procedures were identical to those used in Experiment 2, with the exception that half the participants received two blocks of the ISR-free task and the other half received two blocks of the standard ISR task. Participants in the ISR-free task spoke their responses out loud, so that their output order could be recorded; but due to a methodological oversight, participants in the standard ISR task wrote down their responses in forward serial order but they did not speak their responses out loud.

Results

Analyses of serial position curves of all data. Figure 10 shows the proportion of words recalled in the correct serial position for the standard ISR task (see Figure 10A) and the ISR-free task (see Figure 10B) at each of the 11 different list lengths. A 2 (task) × 11 (list length) mixed ANOVA was performed on the mean ISR scores for the two groups at each list length. This revealed a nonsignificant main effect of task, $F(1, 38) = 2.80$, $MSE = .033$, $p > .05$, a significant main effect of list length, $F(10, 380) = 666.5$, $MSE = .007$, $p < .001$, and a significant two-way interaction, $F(10, 380) = 2.35$, $MSE = .007$, $p < .05$. The two-way interaction occurred because at List Lengths 2–4 there were small but nonsignificant advantages for participants performing standard ISR compared with ISR-free, but the pattern was reversed at list lengths greater than 4, and at list lengths of 10 and greater, there was a small but significant increase in performance in ISR-free relative to standard ISR.

The increase in ISR-free performance at longer list lengths was also observed using FR scoring. A 2 (task) × 11 (list length) mixed ANOVA using FR scoring revealed a nonsignificant main effect of task, $F(1, 38) = 2.27$, $MSE = .032$, $p > .05$, a highly significant main effect of list length, $F(10, 380) = 899.2$, $MSE = .004$, $p < .001$, and a significant interaction, $F(10, 380) = 3.38$, $MSE = .004$, $p < .001$. Pairwise comparisons revealed that participants performing ISR-free recalled almost the same proportion of words

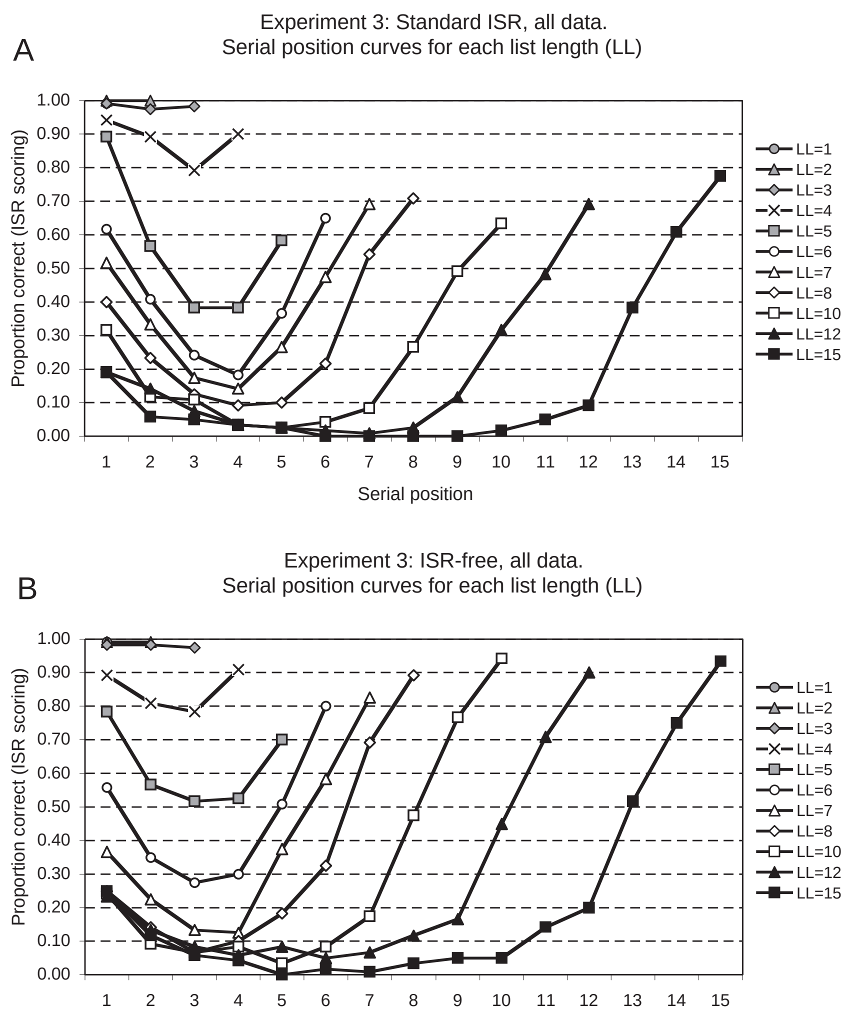


Figure 10. Data from Experiment 3 showing the serial position curves for (A) the standard ISR task and (B) the ISR-free task for each list length. ISR = immediate serial recall.

as those performing standard ISR at List Lengths 1–4, but there was a small increase in performance in ISR-free relative to standard ISR at list lengths greater than 4, which reached significance at list lengths of 10 and greater.

We examined serial position curves on the two tasks using ISR scoring at three representative list lengths: 4, 8, and 15. A 2 (task) × 4 (serial position) mixed ANOVA revealed a nonsignificant main effect of task, $F(1, 38) = 0.94$, $MSE = .047$, $p > .05$, a significant main effect of serial position, $F(3, 114) = 7.15$, $MSE = .019$, $p > .05$, and a nonsignificant interaction, $F(3, 114) = 0.88$, $MSE = .019$, $p > .05$. For both tasks, there were significant primacy and recency effects at List Length 4, and there was little difference in the shapes of the serial position curves of the two tasks.

A 2 (task) × 8 (serial position) mixed ANOVA revealed a nonsignificant main effect of task, $F(1, 38) = 0.83$, $MSE = .082$, $p > .05$, a significant main effect of serial position, $F(7, 266) = 65.81$, $MSE = .041$, $p < .001$, and a significant interaction, $F(7, 266) = 3.60$, $MSE = .041$, $p < .01$. There were significant primacy and recency effects in both tasks, with greater recency than primacy. Inspection of the interaction revealed that there were small but nonsignificant recall advantages in favor of the standard ISR task at Serial Positions 1–4, small but nonsignificant recall advantages in favor of the ISR-free task at Serial Positions 5–7, and a significant recall advantage for the ISR-free task at Serial Position 8.

A 2 (task) × 15 (serial position) mixed ANOVA revealed a significant main effect of task, $F(1, 38) = 8.14$, $MSE = .065$, $p < .01$, a significant main effect of serial position, $F(14, 532) = 128.97$, $MSE = .022$, $p < .001$, and a nonsignificant interaction, $F(14, 532) = 1.42$, $MSE = .022$, $p > .05$. There were significant primacy and recency effects in both tasks, with greater recency than primacy, and greater recall performance overall with ISR-free than standard ISR.

PFR data. Second, we examined which words from a list were the first to be recalled on each trial. Table 4 shows the PFR data for lists of different lengths in the standard ISR and ISR-free tasks. For both tasks, the serial position refers to the input position of the word that was first recalled (rather than the serial position to which that word was assigned). At short list lengths, participants tended to initiate their recall with the first words in the list (those presented at Serial Position 1), but as the list length increased, so this tendency decreased for both tasks, albeit more rapidly with ISR-free than standard ISR. In the ISR-free task, as the list length increased, so there was an increase in initiating recall with one of the last four words in the list. In the standard ISR task, as the list length increased, there was an increase in initiating recall with one of the last four words in the list, and also with one of the earlier serial positions.

Figures 11A (standard ISR) and 11B (ISR-free) summarize these data. A 2 (task: ISR-free and standard ISR) × 10 (list lengths: 2–8, 10, 12, and 15) mixed ANOVA was performed on the proportion of trials starting with Serial Position 1. This revealed a significant main effect of task, $F(1, 38) = 25.54$, $MSE = .120$, $p < .001$, a significant main effect of list length, $F(9, 342) = 223.2$, $MSE = .024$, $p < .001$, and a significant interaction, $F(9, 342) = 6.44$, $MSE = .024$, $p < .001$.

Figures 11A and 11B also show that at longer list lengths there is an increased tendency to initiate recall with words from one of the last four serial positions, and that this tendency increases more rapidly with ISR-free than standard ISR. This pattern was also confirmed in a 2 (task: ISR-free and standard ISR) × 10 (list lengths: 2–8, 10, 12, and 15) mixed ANOVA performed on the proportion of trials starting with one of the last four serial positions. The means are provided later in Table 6. This revealed a significant main effect of task, $F(1, 38) = 33.21$, $MSE = .181$, $p < .001$, a significant main effect of list length, $F(9, 342) = 100.6$, $MSE = .029$, $p < .001$, and a significant interaction, $F(9, 342) = 17.40$, $MSE = .029$, $p < .001$.

Unlike previous experiments, there was also a sizeable minority of trials using standard ISR in which participants started their recall with serial positions other than the first item or one of the last four words (see the “other” data in Figure 11B). We examined whether the tendency to start with an early list item (observed with Serial Position 1) extended to Serial Positions 2, 3, or 4. We compared the mean proportion of trials in which words presented in Serial Positions 2, 3, and 4 were the first word recalled at each list length from 5 to 15. The mean proportions of words first recalled from Serial Positions 2, 3, and 4 were .016, .029, and .035, respectively, for ISR-free, but were .068, .041, and .041 for Serial Positions 2, 3, and 4, respectively, for standard ISR. A three-way mixed ANOVA revealed that the proportion of trials starting with one of these early serial positions was far greater in standard ISR than ISR-free, $F(1, 38) = 15.96$, $MSE = .007$, $p < .001$, and decreased with increasing list length, $F(6, 228) = 9.78$, $MSE = .004$, $p < .001$. Critically, there was a significant interaction between task and serial position, $F(2, 76) = 7.89$, $MSE = .006$, $p < .001$, which when analyzed showed that there was a heightened tendency to start recall with Serial Position 2 relative to Serial Positions 3 and 4 for the standard ISR task, but no such tendency was present in the ISR-free data. Finally, there was also a significant interaction between serial position and list length, $F(12, 456) = 2.68$, $MSE = .007$, $p < .01$. The remaining main effect, two-way interaction, and three-way interaction were not significant (all Fs < 1). Thus, there was a general tendency to recall more early list items in the standard ISR task compared to ISR-free (reflecting an increase in the “other” first responses in Figure 11B), there was greater first recall of early list items more with shorter lists, and there was some evidence for a graded primacy advantage with standard ISR but not ISR-free.

The effect of first recall on the serial position curves. We split the serial position curve data by whether participants initiated recall with the first word or one of the last four words. Figure 12A shows the serial position curves for the standard ISR trials in which the first word recalled was from Serial Position 1; Figure 12B shows the same data plotted by FR scoring; and Figure 12C shows the serial position curves for the standard ISR trials in which the first word recalled was from one of the last four serial positions. As in earlier studies, the shapes of the serial position curves in Figures 12A, 12B, and 12C are different: There is an increased tendency for heightened primacy with reduced recency when the first word is recalled first in Figure 12A, elevated levels of recall on early items in Figure 12B, but extended recency with reduced primacy when one of the last four words is first recalled in Figure 12C.

Figures 12D, 12E, and 12F show the equivalent data split for the ISR-free task. Replicating the ISR-free data of Experiment 2 (see Figures 8D, 8E, and 8F), there is an increased tendency for

Table 4
Data From Experiment 3 for Standard ISR and ISR-Free

Serial position	List length 1	2	3	4	5	6	7	8	10	12	15
						Standard ISR					
1	*119*	*120*	*120*	*113*	*108*	*81*	*68*	*54*	*44*	*25*	*23*
2		**0**	**0**	**4**	**4**	15	12	6	7	5	8
3			**0**	**3**	**6**	**7**	5	6	2	5	3
4				**0**	**2**	**9**	**11**	8	1	2	1
5					**0**	**8**	**17**	**17**	10	8	3
6						**0**	**6**	**13**	10	2	1
7							**1**	**13**	**12**	7	2
8								**3**	**18**	7	3
9									**11**	**20**	4
10									**5**	**22**	3
11										**14**	12
12										**2**	**14**
13											**29**
14											**12**
15											**2**
Void	1	0	0	0	0	0	0	0	0	1	0
Error on first word	0	0	0	0	0	0	0	0	0	0	0
Total	120	120	120	120	120	120	120	120	120	120	120
						ISR-free					
1	*119*	*117*	*118*	*105*	*93*	*55*	*27*	*12*	*8*	*4*	*2*
2		**1**	**2**	**8**	**2**	5	4	1	0	1	0
3			**0**	**3**	**10**	**7**	2	4	1	0	0
4				**3**	**0**	**11**	**7**	10	0	1	0
5					**15**	**17**	**24**	**9**	2	0	0
6						**24**	**21**	**19**	3	0	0
7							**32**	**22**	**12**	2	0
8								**43**	**24**	5	0
9									**24**	**13**	0
10									**46**	**23**	3
11										**16**	5
12										**54**	**11**
13											**26**
14											**13**
15											**60**
Void	1	2	0	1	0	0	0	0	0	0	0
Error on first word	0	0	0	0	0	1	3	0	0	1	0
Total	120	120	120	120	120	120	120	120	120	120	120

Note. The data show the distribution of the first words recalled on each trial, as a function of the list length and the words' serial position. The italicized values represent the frequency of trials in which the first word recalled was from Serial Position 1, and the bold values represent the frequency of trials in which the first word recalled was from one of the last four serial positions. The values in regular font represent the frequency of trials in which the first word recalled was from one of the other serial positions. ISR = immediate serial recall; void = no words were recalled on a particular trial.

heightened primacy with reduced recency when the first word is recalled first in Figure 12D, elevated recall in Figure 12E, but extended recency with reduced primacy when one of the last four words is first recalled in Figure 12E.

One striking feature of these data is the similarity between the two tasks when the first response is controlled for. The data at List Length 6 were examined across the two tasks using ISR-scoring on the data in Figures 12A and 12D in which recall initiated with Serial Position 1. A total of 17 participants from the ISR-free task (*M*s = .90, .46, .25, .20, .40, and .70) and all 20 participants from the standard ISR task (*M*s = .91, .45, .28, .19, .35, and .64) contributed to this analysis. Ignoring Serial Position 1, a 2 (task) × 5 (serial position: 2–5) mixed ANOVA revealed a nonsignificant main effect of task, $F(1, 35) = 0.16$, $MSE = .134$, $p > .05$, a significant main effect of serial position, $F(4, 140) = 17.02$, $MSE = .073$, $p < .0001$, and a nonsignificant interaction, $F(4, 140) = 0.13$, $MSE = .073$, $p > .05$. Thus, once equated for starting with Serial Position 1, there was similar performance on both tasks. Pairwise comparisons revealed that there were primacy and recency effects.

When the same List Length 6 data were compared using FR scoring (see Figures 12B and 12E) over Serial Positions 2–6, there were highly similar findings (means for standard ISR were 1.00, .52, .58, .44, .63, and .77; for ISR-free, the means were 1.00, .53, .40, .50, .62, and .82). The ANOVA revealed that there was no significant main effect of task, $F(1, 35) = 0.16$, $MSE = .134$, $p > .05$, a significant main effect of serial position, $F(4, 140) = 17.02$, $MSE = .073$, $p < .0001$, and a nonsignificant interaction, *F*(4,

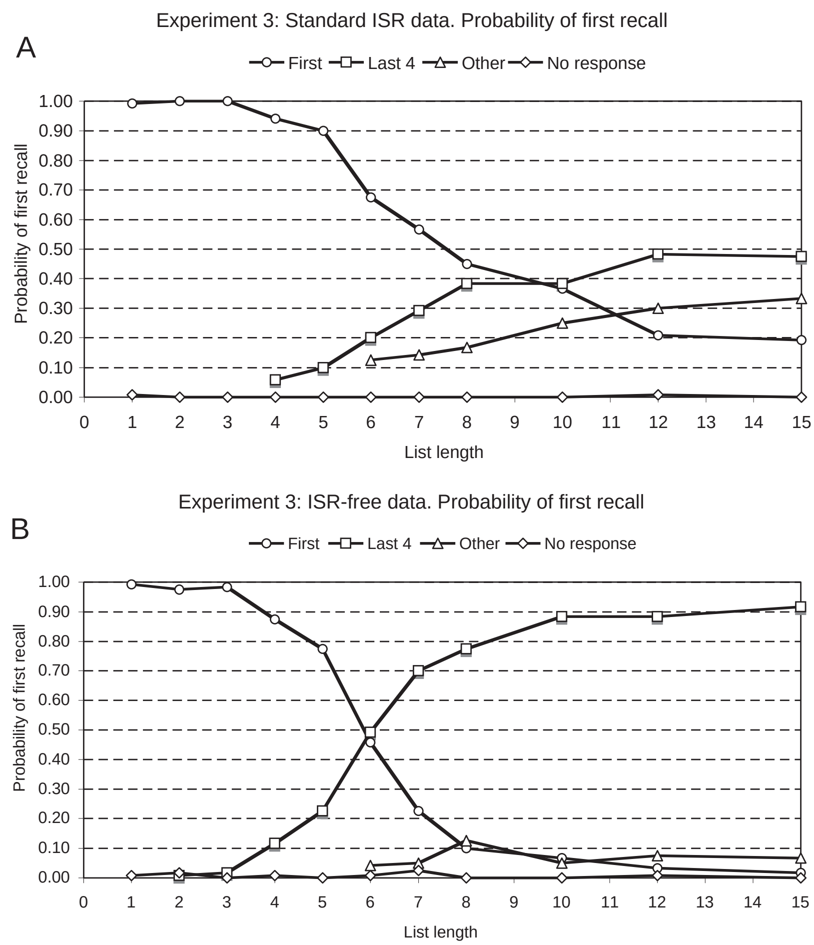


Figure 11. Data from the standard immediate serial recall (ISR) task (A) and the ISR-free task (B) of Experiment 3 showing the proportion of trials at each list length in which recall initiated with the first word in the list, one of the last four words in the list, or one of the other words in the list. On a small minority of trials, participants either recalled no words from the list (void) or began recall with a word not on the list (error).

140) = 0.14, *MSE* = .073, $p > .05$. Pairwise comparisons again revealed significant primacy and recency effects.

Finally, an analysis was performed on trials starting with one of the last four serial positions using the List Length 6 data and ISR scoring (see Figures 12C and 12F) over Serial Positions 1–6. Nineteen participants from the ISR-free group (*M*s = .32, .27, .53, .66, .81, and 1.00) and 13 participants from the standard ISR group (*M*s = .08, .23, .41, .60, .68, and .96) contributed data to this analysis. There was a significant main effect of task, $F(1, 30) = 5.22$, *MSE* = .180, $p < .05$, a significant main effect of serial position, $F(5, 150) = 40.94$, *MSE* = .074, $p < .001$, and a nonsignificant interaction, $F(5, 150) = 1.28$, *MSE* = .074, $p > .05$. Both tasks showed significant extended recency effects, but performance was superior overall for ISR-free.

Analyses of output order using CRPs. Finally, we examined the extent to which the output orders showed evidence of forward-ordered recall. For both tasks, the lag refers to the difference in serial positions based on the input positions of the words that were recalled (rather than the serial positions to which the words were assigned). Figure 13 plots the CRP values for lag + 1 responses as a function of list length for standard ISR and ISR-free. The CRP values for lag + 1 responses were analyzed by a 2 (task: standard ISR and ISR-free) × 10 (list length: 2, 3, 4, 5, 6, 7, 8, 10, 12, and 15) mixed ANOVA, which revealed a nonsignificant main effect of task, $F(1, 38) = 0.06$, *MSE* = .096, $p > .05$, a significant main effect of list length, $F(9, 342) = 64.06$, *MSE* = .021, $p < .001$, and a nonsignificant interaction, $F(9, 342) = 1.41$, *MSE* = .021, $p > .05$. Thus there is a similar tendency for forward-ordered recall to decrease with increasing list length for both tasks, with a similar degree of forward-ordered recall on both tasks. Note that even at the longer list lengths there remains considerable forward-ordered recall (CRP values for lag + 1 remain greater than .50) for both tasks, even though there is no formal requirement to output in forward order in ISR-free.

Discussion

Considering first the ISR-free data, Experiment 3 essentially replicated the ISR-free data from Experiment 2. When the list was short, it was common to begin ISR-free with the first list item

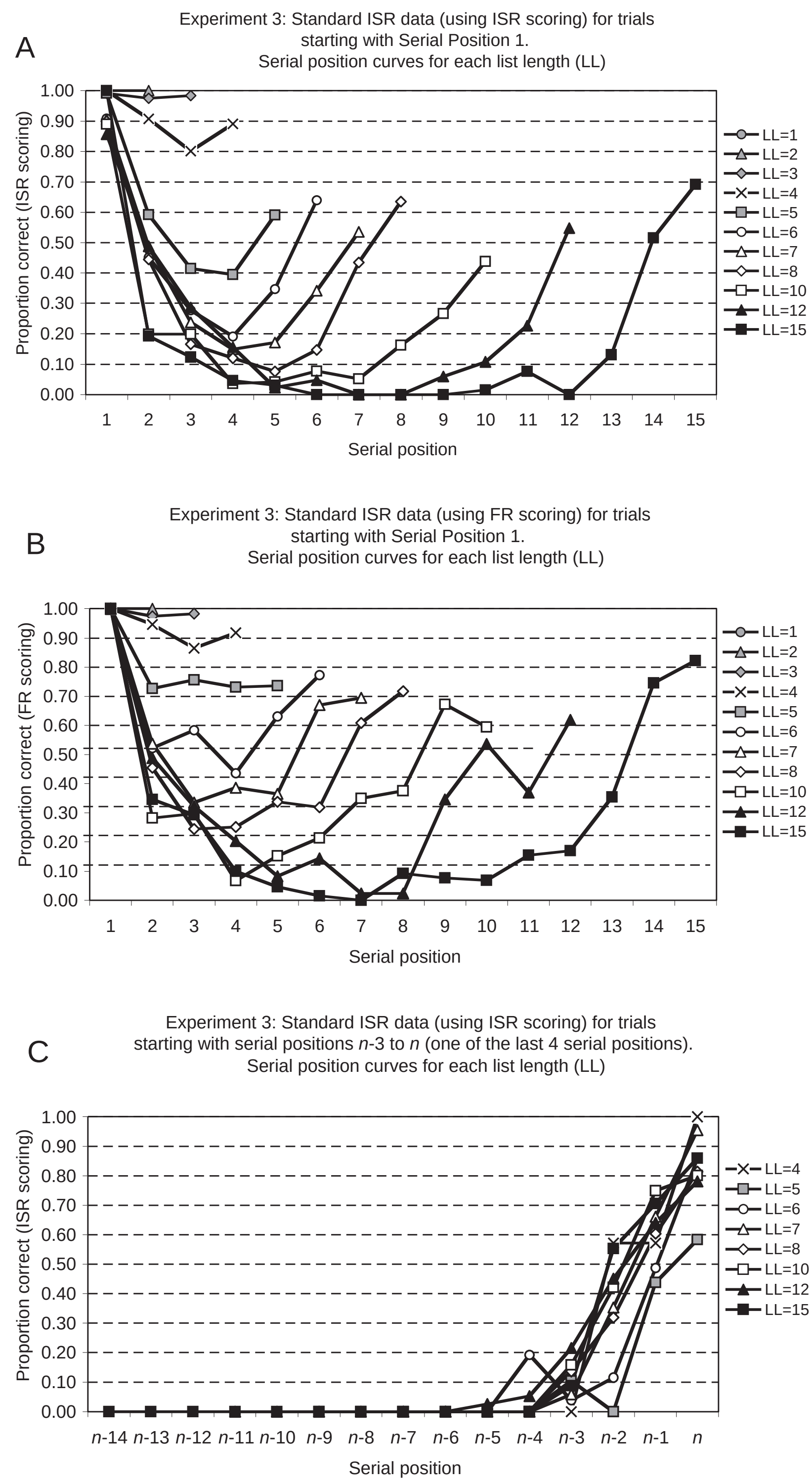
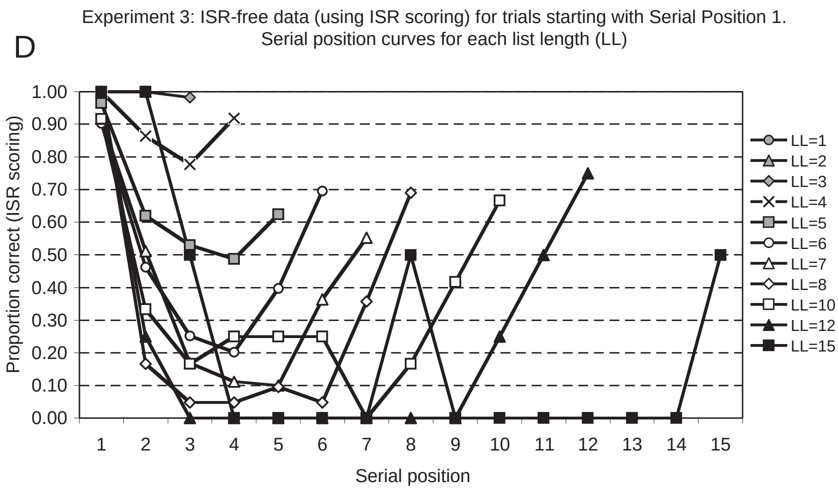
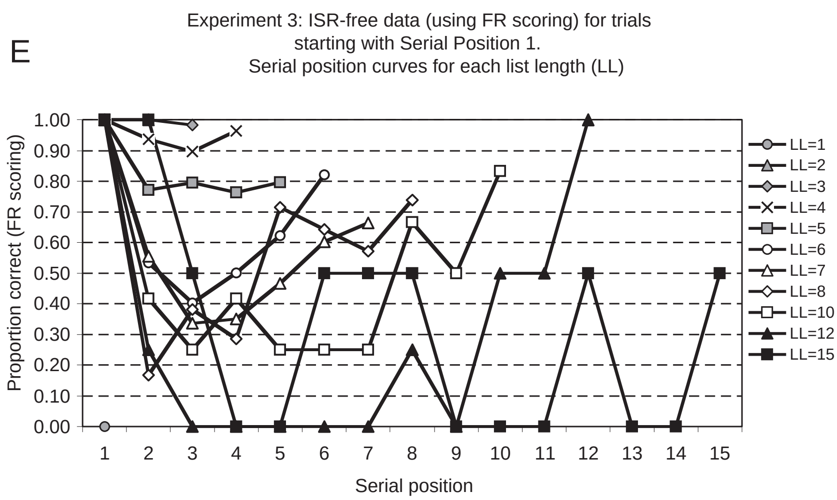


Figure 12. Data from Experiment 3 showing serial position curves for standard immediate serial recall (ISR) trials that began with Serial Position 1 using (A) ISR scoring and (B) free recall (FR) scoring and (C) for trials that began with one of the last four serial positions using FR scoring. Panels D–F show serial position curves for the ISR-free data for trials which began with Serial Position 1 using (D) ISR scoring and (E) FR scoring and (F) for trials that began with one of the last four serial positions using ISR scoring. In Panels C and F, the serial positions have been recency justified.





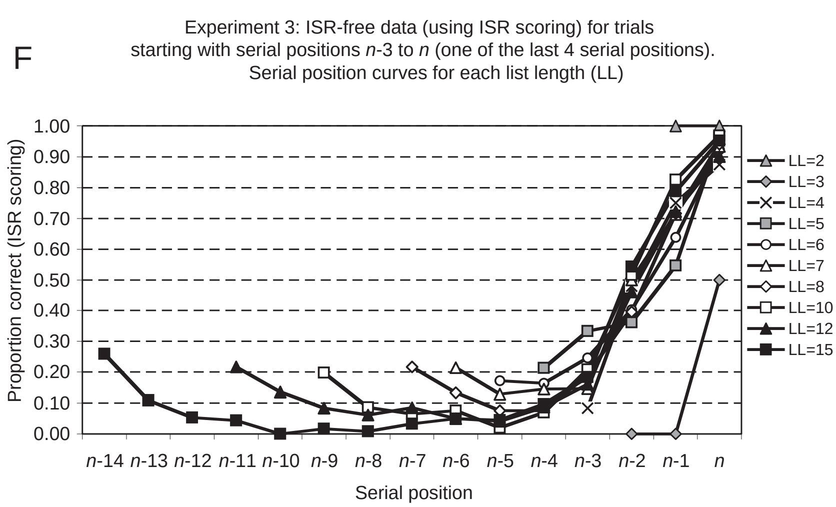


Figure 12. (continued)

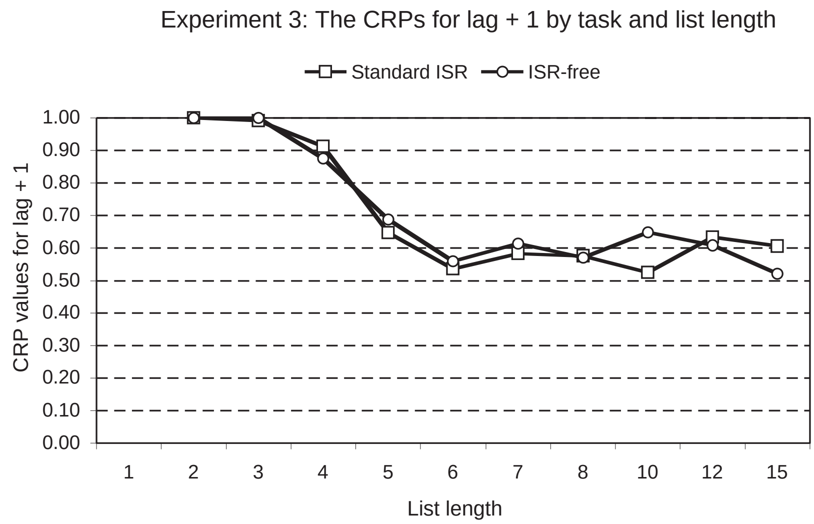


Figure 13. Data from Experiment 3 showing the proportion of lag + 1 responses as a function of list length and task. CRP = conditionalized response probability.

(even though participants did not have to do so), but as the list length was increased so there was an increased tendency for ISR-free to begin with one of the last four list items. The first word recalled affected the resultant serial position curves. Enhanced recall of early list items was found when recall started with Serial Position 1, but extended recency effects were observed when recall started with one of the last four items.

Considering next the standard ISR data, Experiment 3 showed that there was a similar tendency to recall Serial Position 1 first for short lists, and, as the list length increased, so participants showed a large reduction in recalling the first word first (although this reduction was attenuated relative to ISR-free). When participants could not recall the first item in the standard ISR task, they started with an earlier ("other") item more often in the standard ISR task than in the ISR-free task. Moreover, there was some evidence for a graded primacy effect in the PFR in the standard ISR task: There was a greater proportion of trials starting with Serial Position 2 than 3 or 4. However, there was still a sizeable minority of trials in which recall initiated with one of the last four serial positions. These data are, of course, consistent with participants observing the standard ISR task instructions, and they are furthermore consistent with participants possessing an amount of strategic control over which words to output first (for related data, see also Bhatarah et al., 2008; Lewandowsky et al., 2008; Tan & Ward, 2007).

Interestingly, the effects of start position on the resultant serial position curves are similar on the two tasks: Greater primacy is found when recall starts with the first word in the list, but greater recency is found when recall was from one of the last four words. Moreover, there were similar degrees of forward-ordered recall in both tasks, even though forward-ordered recall was not a necessary requirement of the ISR-free task. These similarities in the degree of forward-ordered recall suggest that ISR-free is best considered a variant of ISR, rather than IFR.

Experiment 4

Experiment 3 compared ISR with ISR-free and showed that in both tasks there was a reduction in initiating recall with the first list item as the length of the list increased. Experiment 4 examined whether a similar pattern of data could be observed using different variants of the RoO task (for a recent review, see Lewandowsky et al., 2008). In RoO tasks, participants are presented with sequences of words for immediate recall as in the previous experiments, but at test all the list items are re-presented in a new random order and participants must place each word into the correct serial position. The task is therefore well suited for examining the effects of output order and list length.

In Experiment 4, we examined performance on the standard version of the RoO task, in which participants must reconstruct the order of the re-presented list items in a strict forward order, by allocating words into their locations from top to bottom in the response grid. We compared performance on this condition with free RoO, in which participants were free to allocate the re-presented words into their positions in the grid in whatever order they liked. Through the positioning and order of responses, the task allows the recording of both serial position and output order without the need to ask participants to speak out loud during one or both conditions.

Method

Participants. Thirty-four volunteers from the University of Essex participated in this experiment. None had participated in any of the earlier three experiments.

Materials and apparatus. The materials were the same as those used in Experiment 3, except that participants were not provided with a booklet of response grids.

Design. The experiment used a mixed design. The type of task was manipulated between subjects with two levels: standard RoO and free RoO. There were also two within-subjects independent variables: list length, with 11 levels (1, 2, 3, 4, 5, 6, 7, 8, 10, 12, and 15), and serial position, with up to 15 levels. The dependent variable was the proportion of words recalled in the correct serial position.

Procedure. The method of presentation of the stimuli during study was identical to that used in Experiments 2 and 3. However, the method of testing differed. After the last word had been presented, the participants were presented with all the list items arranged in a new random order in a column on the left-hand side of the screen, and a response grid of the same type as presented in Experiments 2 and 3 was presented on the right-hand side of the screen. In the first column of the response grid in ascending order were the numbers 1 to list length. The second column of the response grid was left blank. Unlike the recall experiments which involved writing down the words, in Experiment 4, the participants were required to fill the response grid by first clicking on the word to be entered (from the column on the left) and then clicking on the location in the response grid corresponding to that word's list position on the right.

The computer program in the standard RoO task only allowed participants to fill the response grid from top to bottom, that is, in strict serial order. By contrast, participants in the free RoO task could click on any word and click on any location. It was not possible to change the responses once they had been entered. When the response grid was completed, participants were free to press a computer button on the screen to continue.

Results

Analyses of serial position curves of all data. Figure 14 shows the proportion of words recalled in the correct serial position for the standard RoO task (see Figure 14A) and the free RoO task (see Figure 14B) at each of the 11 different list lengths. The graphs exclude the data from a small minority of trials (28 trials out of a total of 2,244, less than 1.3%) in which participants clicked on a word followed by two locations in succession, resulting in the same item being placed in two different locations. A 2 (task) × 11 (list length) mixed ANOVA was performed on the mean RoO scores for the two groups at each list length. This revealed a significant main effect of task, $F(1, 32) = 18.51$, $MSE = .042$, $p < .001$, a significant main effect of list length, $F(10, 320) = 381.1$, $MSE = .008$, $p < .001$, and a significant two-way interaction, $F(10, 320) = 7.14$, $MSE = .008$, $p < .001$. The two-way interaction occurred because at List Lengths 1–4 there were no significant differences in levels of recall for the two tasks, but at list lengths above 5 there were significant reconstruction advantages for the free RoO conditions over the standard RoO conditions.

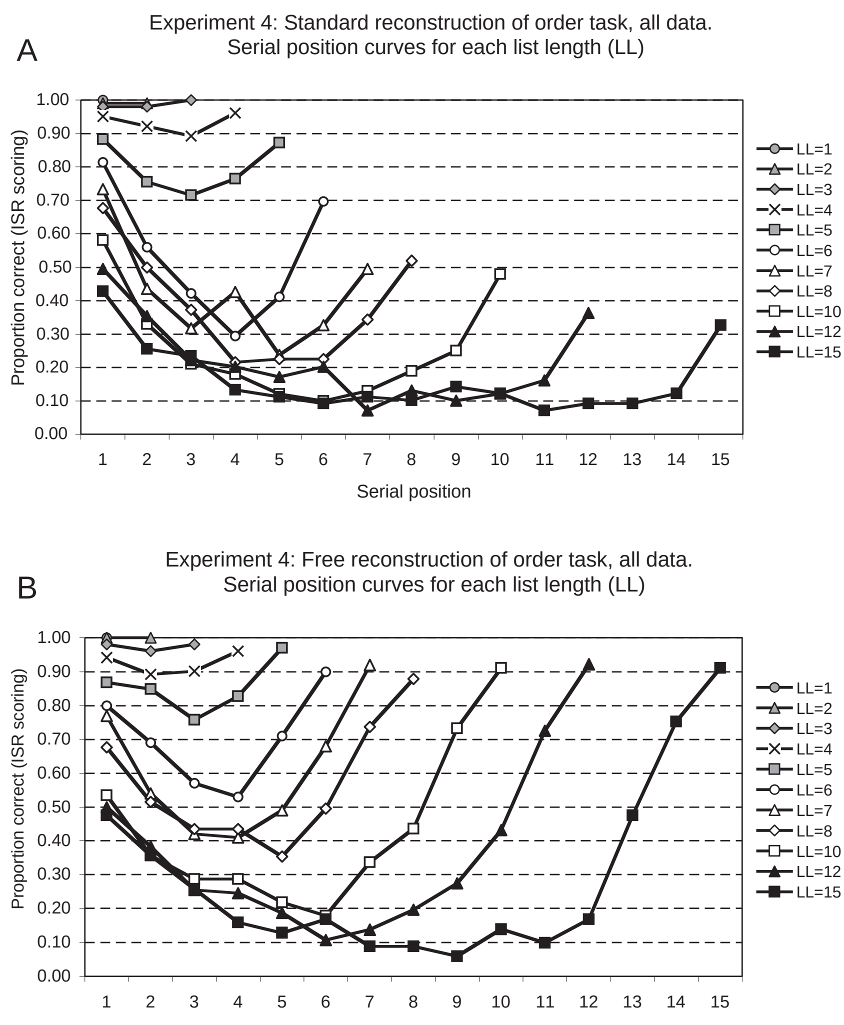


Figure 14. Data from Experiment 4 showing the serial position curves for (A) the standard reconstruction of order (RoO) task and (B) the free RoO task for each list length. ISR = immediate serial recall.

A 2 (task) × 4 (serial position) mixed ANOVA on the serial position curve at List Length 4 revealed a nonsignificant main effect of task, $F(1, 32) = 0.51$, $MSE = .039$, $p > .05$, a significant main effect of serial position, $F(3, 96) = 4.52$, $MSE = .007$, $p < .01$, and a nonsignificant interaction, $F(3, 96) = 0.32$, $MSE = .007$, $p > .05$. For both tasks there were significant primacy and recency effects at List Length 4, and there was little difference in the shapes of the serial position curves of the two tasks.

A 2 (task) × 8 (serial position) mixed ANOVA revealed a significant main effect of task, $F(1, 32) = 14.31$, $MSE = .151$, $p < .001$, a significant main effect of serial position, $F(7, 224) = 25.93$, $MSE = .032$, $p < .001$, and a significant interaction, $F(7, 224) = 5.84$, $MSE = .032$, $p < .01$. There were significant primacy and recency effects with both tasks. However, although there was near equivalent performance on Serial Positions 1–3 on the two tasks, there were significant recall advantages at Serial Positions 4, 6, 7, and 8 for the free RoO task relative to the standard RoO task.

A 2 (task) × 15 (serial position) mixed ANOVA revealed a significant main effect of task, $F(1, 32) = 15.55$, $MSE = .132$, $p < .001$, a significant main effect of serial position, $F(14, 448) = 29.81$, $MSE = .030$, $p < .001$, and a significant interaction, $F(14, 448) = 13.80$, $MSE = .030$, $p < .001$. There were significant primacy and recency effects on both tasks, but whereas there were similar levels of primacy for the two tasks, there was significantly greater recency on the last three serial positions for the free RoO task.

PFR data. Second, we examined which words from a list were the first to be recalled on each trial. Table 5 shows the PFR data for lists of different lengths in the standard RoO and free RoO tasks. For both tasks, the serial position refers to the input position

Table 5
Data From Experiment 4 for Standard RoO and Free RoO

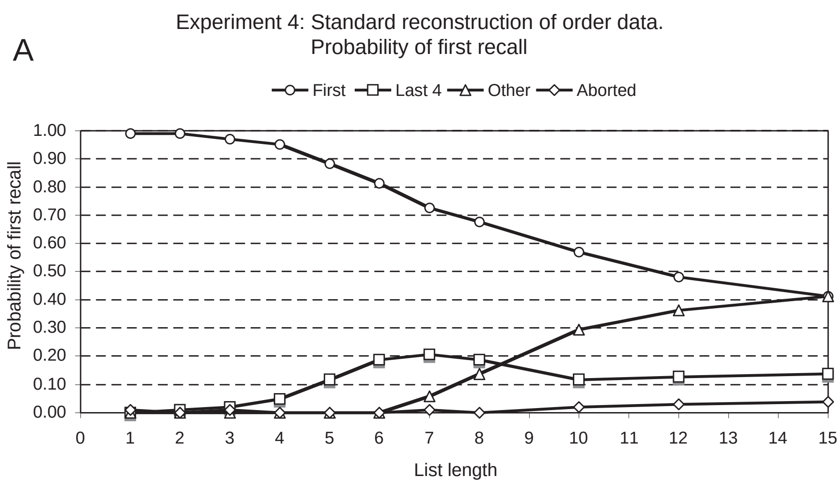
	List length										
Serial position	1	2	3	4	5	6	7	8	10	12	15
						Standard RoO					
1	*101*	*101*	*99*	*97*	*90*	*83*	*74*	*69*	*58*	*49*	*42*
2		**1**	**2**	**2**	**7**	8	6	8	9	4	6
3			**0**	**3**	**3**	**7**	6	6	4	4	6
4				**0**	**2**	**3**	**7**	6	9	5	4
5					**0**	**1**	**6**	**8**	8	6	8
6						**0**	**2**	**2**	4	7	4
7							**0**	**3**	**3**	11	1
8								**0**	**3**	5	2
9									**1**	**2**	7
10									**1**	**4**	4
11										**1**	3
12										**1**	**5**
13											**3**
14											**2**
15											**1**
Aborted	1	0	1	0	0	0	1	0	2	3	4
Total	102	102	102	102	102	102	102	102	102	102	102
						Free RoO					
1	*102*	*100*	*95*	*91*	*70*	*55*	*46*	*21*	*17*	*12*	*11*
2		**0**	**3**	**2**	**2**	3	2	1	3	1	0
3			**2**	**4**	**8**	**3**	2	1	1	0	0
4				**5**	**10**	**6**	**3**	2	1	1	0
5					**9**	**23**	**15**	**13**	4	1	0
6						**10**	**13**	**11**	2	2	0
7							**19**	**22**	**7**	4	0
8								**28**	**11**	0	0
9									**32**	**13**	1
10									**23**	**9**	0
11										**28**	0
12										**31**	**3**
13											**23**
14											**24**
15											**39**
Aborted	0	2	2	0	3	2	2	3	1	0	1
Total	102	102	102	102	102	102	102	102	102	102	102

Note. The data show the distribution of the first words recalled on each trial, as a function of the list length and the words' serial position. The italicized values represent the frequency of trials in which the first word recalled was from Serial Position 1, and the bold values represent the frequency of trials in which the first word recalled was from one of the last four serial positions. The values in regular font represent the frequency of trials in which the first word recalled was from one of the other serial positions. RoO = reconstruction of order; aborted = trial aborted due to participant selecting multiple locations for the same word.

of the word that was first recalled (rather than the serial position to which that word was assigned). In the vast majority of all trials with short list lengths, recall started with the first word in the list. As the list length increased, so this tendency decreased for both tasks, but it decreased more rapidly with free RoO than with standard RoO.

Figures 15A (standard RoO) and 15B (free RoO) summarize these PFR data. A 2 (task: standard RoO and free RoO) × 10 (list lengths: 2–8, 10, 12, and 15) mixed ANOVA was performed on the proportion of trials starting with Serial Position 1. This revealed a significant main effect of task, $F(1, 32) = 30.22$, $MSE = .159$, $p < .001$, a significant main effect of list length, $F(9, 288) = 79.97$, $MSE = .033$, $p < .001$, and a significant interaction, $F(9, 288) = 7.21$, $MSE = .033$, $p < .001$. There was no significant difference in the proportion of trials starting with Serial Position 1 at List Lengths 2–4, but from List Length 5 there was a significant decrease in the proportion of trials starting with Serial Position 1 relative to the standard RoO task.

Figures 15A and 15B also show that at longer list lengths there is an increased tendency to initiate recall with words from one of the last four serial positions (especially for the free RoO task) and from other serial positions (especially for the standard RoO task). We examined whether the tendency to start with an early list item (observed with Serial Position 1) extended to Serial Positions 2, 3, and 4. We compared the mean proportion of trials in which words presented in Serial Positions 2, 3, and 4 were the first word recalled at each list length from 5 to 15. The mean proportions of words first recalled from Serial Positions 2, 3, and 4 were .017, .021, and .032, respectively, for free RoO, and .068, .050, and .050 for Serial Positions 2, 3, and 4, respectively, for standard RoO. A three-way mixed ANOVA revealed that the proportion of trials starting with one of these early serial positions was far greater in the standard RoO task than the free RoO task, $F(1, 32) = 14.25$, $MSE = .014$, $p < .001$. However, the only other significant main effect or interaction was for the two-way interaction between task and list length, $F(6, 192) = 3.38$, $MSE = .006$, $p < .01$, which when analyzed showed that the tendency to start recall with one of Serial Positions 2, 3, or 4 decreased with increasing list length for the free RoO but remained relatively constant for the standard RoO task. The remaining main effects, two-way interactions, and three-



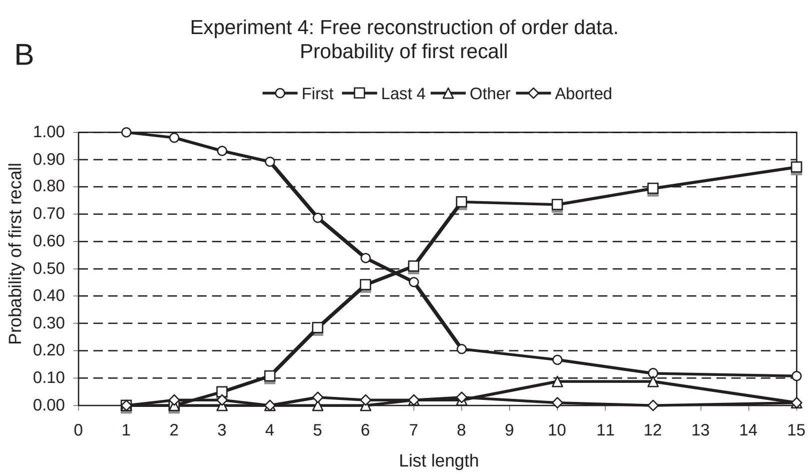


Figure 15. Data from (A) the standard reconstruction of order (RoO) task and (B) the free RoO task of Experiment 4 showing the proportion of trials at each list length in which recall initiated with the first word in the list, one of the last four words in the list, or one of the other words in the list. On a small minority of trials, participants selected multiple locations for the same word and the trial was aborted.

way interaction were not significant. Thus, there was a general tendency to recall more early list items in the standard RoO task compared to the free RoO task, but there was no evidence for a graded primacy advantage with either task.

The effect of first recall on the serial position curves. Figure 16A shows the serial position curves for the free RoO trials in which the first word recalled was from Serial Position 1, and Figure 16B shows the serial position curves for the free RoO trials in which the first word recalled was from one of the last four serial positions. As in earlier studies, the shapes of the serial position curves in Figures 16A and 16B are rather different: There is an increased tendency for heightened primacy when the first word is recalled first in Figure 16A, but more extended recency with reduced primacy in Figure 16B when one of the other words is recalled first.

For List Length 6, the means for 17 participants in the free RoO task who initiated at least one trial with Serial Position 1 were .97, .77, .66, .45, .60, and .81 (ISR scoring), and the means for the 17 participants in the standard RoO task who initiated at least one trial with Serial Position 1 were 1.00, .64, .48, .33, .42, and .69 (ISR scoring). Ignoring Serial Position 1, a 2 (task) × 5 (serial position: 2–6) mixed ANOVA revealed a main effect of task that approached significance, $F(1, 32) = 4.07$, $MSE = .215$, $p < .06$, a significant main effect of serial position, $F(4, 128) = 12.88$, $MSE = .055$, $p < .0001$, and a nonsignificant interaction, $F(4, 128) = 0.16$, $MSE = .055$, $p > .05$. Pairwise comparisons revealed significant primacy and recency effects.

The means for 15 participants in the free RoO task who initiated at least one trial with one of the last four serial positions were .59, .63, .40, .49, .73, and .98 (ISR scoring), which differed significantly, $F(5, 70) = 8.08$, $MSE = .078$, $p < .001$. Pairwise comparisons revealed significant recency effects and nonsignificant primacy effects.

Analyses of output order using CRPs. Finally, we examined the extent to which the output orders showed evidence of forward-ordered recall. For both tasks, the lag refers to the difference in

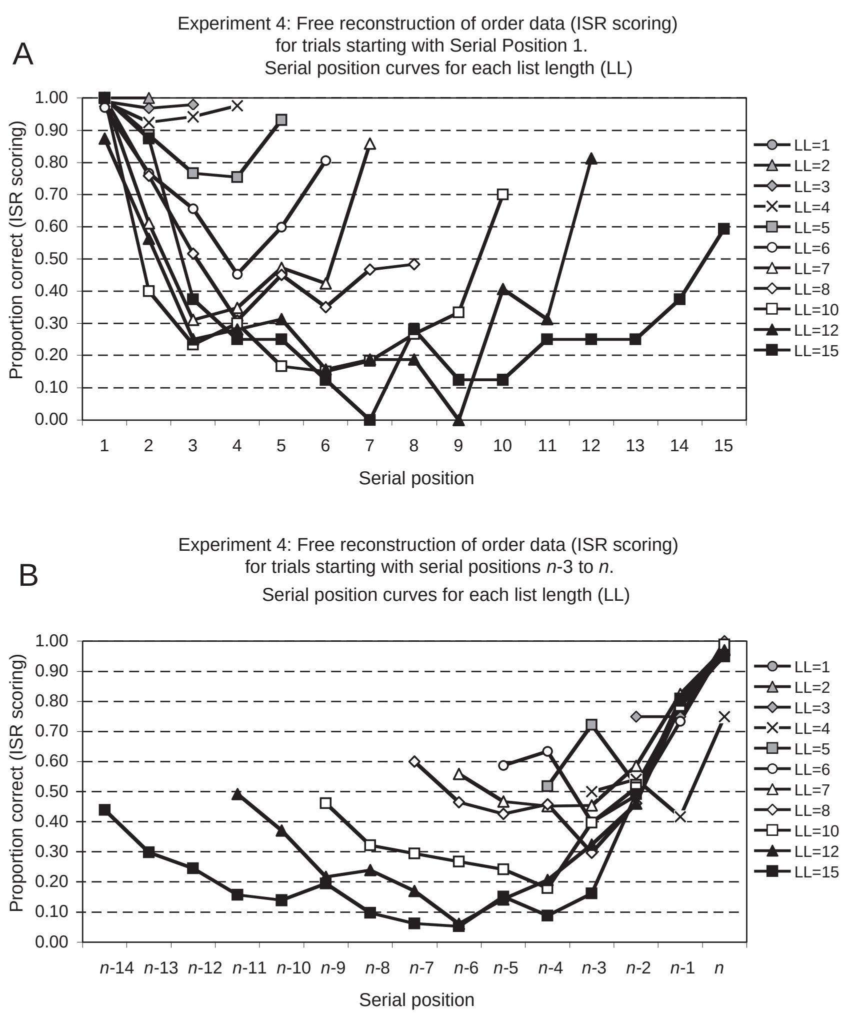


Figure 16. Data from Experiment 4 showing the serial position curves for the free reconstruction of order (RoO) task (A) for trials which began with Serial Position 1 using immediate serial recall (ISR) scoring and (B) for trials which began with one of the last four serial positions using free recall (FR) scoring. In Panel B, the serial positions have been recency justified.

serial positions based on the input positions of the words that were recalled (rather than the serial positions to which the words were assigned). Figure 17 plots the CRP values for lag + 1 responses as a function of list length for standard RoO and free RoO. The CRP values for lag + 1 responses were analyzed by a 2 (task: standard RoO and free RoO) × 10 (list length: 2, 3, 4, 5, 6, 7, 8, 10, 12, and 15) mixed ANOVA, which revealed a significant main effect of task, $F(1, 32) = 14.81$, $MSE = .050$, $p < .001$, a significant main effect of list length, $F(9, 288) = 221.5$, $MSE = .011$, $p < .001$, and a significant interaction, $F(9, 288) = 4.83$, $MSE = .011$, $p < .001$. An analysis of the interaction revealed that there was no significant difference in the CRP values for lag + 1 between the two tasks at List Lengths 2–5, but there was significantly heightened forward-ordered reconstruction in the free RoO task for List Lengths 6 and greater. Thus there is reduced forward output with increasing list lengths for both tasks, but the reduction is steeper in the standard RoO task at longer list lengths.

Discussion

Experiment 4 showed similar findings to Experiment 3 but used RoO tasks rather than serial recall tasks. At short list lengths participants tended to initiate recall with the first word in the list in both the standard RoO task and the free RoO task, and this tendency decreased as the list length increased. As the list length increased, participants tended to start their reconstructions with one of the last four list items in the free RoO task, but in the standard RoO task, participants started with other list items.

We note that there was less of a tendency to start with one of the last four conditions with increasing list length in the standard RoO task compared to the standard ISR task of Experiment 3, but this is likely to reflect participants' opportunity to omit early serial positions in the recall task of Experiment 3, whereas they were compelled to select items for all earlier serial positions before the last serial positions in the standard RoO task.

General Discussion

A summary of the effects of increasing list length in all four experiments is presented in Table 6 (for ease of comparison across experiments, Experiment 1 data from List Lengths 9, 11, 13, and 14 are not presented). Consistent with known list length effects in IFR (e.g., Murdock, 1962; Ward, 2002) and ISR (e.g., Crannell & Parrish, 1957; Drewnowski & Murdock, 1980), the proportion of words correctly recalled decreased with increasing list lengths for all tasks. However, what is most striking and new in our experiments is that, in all four experiments, on all immediate memory tasks, participants tended to start their recall with the first word in the list for short lists, but as the list length was increased, so there was a decreased tendency to start with the first list item. This tendency was found in the standard ISR task (Experiment 3) and the standard RoO task (Experiment 4) where participants were instructed to start their recall with the first item. Critically, this tendency was also shown in the IFR task (Experiments 1–2), the ISR-free task (Experiments 2–3), and the free RoO task (Experiment 4), even though these tasks did not require that the first word in the list be output first.

As the list length increased, so there was an increased tendency to start with one of the last serial positions when the participants were free to do so. Thus, in IFR (Experiments 1–2), ISR-free (Experiments 2–3), and free RoO (Experiment 4), participants increasingly tended to recall one of the most recent four words first at increasing list lengths. This tendency was also observed (to a lesser extent) in standard ISR (Experiment 3), but although possible, such outputs necessarily resulted in all the early and middle list items being omitted. Such a strategy is not possible in the standard RoO task (Experiment 4), where participants were compelled to select the word corresponding to Serial Position 1 first.

In all tasks, there was also a high tendency to recall in forward serial order, a tendency that decreased with increasing list length. The list length and the start position strongly influenced the shape of the

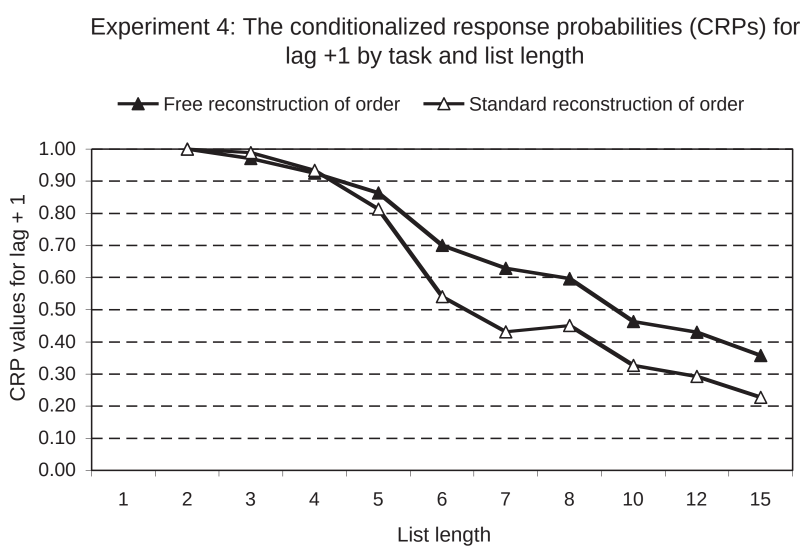


Figure 17. Data from Experiment 4 showing the proportion of lag + 1 responses as a function of list length and task. CRP = conditionalized response probability.

Table 6
Overview of the Main Findings Across the Different Tasks in the Four Experiments

Experiment and task	List length 1	2	3	4	5	6	7	8	10	12	15
A: Overall recall: FR scoring											
Experiment 1											
IFR	.99	.94	.91	.81	.69	.60	.56	.49	.41	.37	.30
Experiment 2											
IFR	1.00	.99	.98	.94	.77	.68	.58	.53	.44	.35	.30
ISR-free	1.00	1.00	.99	.93	.78	.63	.59	.50	.42	.35	.32
Experiment 3											
ISR-free	.99	.99	.99	.93	.79	.63	.54	.49	.41	.36	.32
Standard ISR	.99	1.00	.99	.93	.77	.62	.54	.45	.36	.28	.24
B: Overall recall: ISR scoring											
Experiment 2											
ISR-free	.99	1.00	.99	.89	.64	.46	.41	.33	.27	.21	.18
Experiment 3											
ISR-free	.99	.99	.98	.85	.62	.47	.38	.33	.30	.25	.21
Standard ISR	.99	1.00	.98	.88	.56	.41	.37	.30	.21	.18	.15
Experiment 4											
Free RoO	1.00	1.00	.97	.92	.85	.70	.61	.56	.43	.36	.29
Standard RoO	1.00	.99	.99	.93	.80	.53	.42	.38	.26	.22	.16
C: Probability of first recall = Serial Position 1											
Experiment 1											
IFR	.99	.88	.88	.70	.56	.36	.27	.22	.14	.08	.11
Experiment 2											
IFR	1.00	.98	.97	.89	.73	.58	.35	.31	.19	.07	.05
ISR-free	.99	1.00	.97	.85	.68	.42	.33	.18	.15	.10	.06
Experiment 3											
ISR-free	.99	.98	.98	.88	.78	.46	.23	.10	.07	.03	.02
Standard ISR	.99	1.00	1.00	.94	.90	.68	.57	.45	.37	.21	.19
Experiment 4											
Free RoO	1.00	.98	.93	.89	.69	.54	.45	.21	.17	.12	.11
Standard RoO	.99	.99	.97	.95	.88	.81	.73	.68	.57	.48	.41
D: Probability of first recall = One of the last four serial positions											
Experiment 1											
IFR	.00	.08	.10	.29	.42	.57	.63	.63	.73	.80	.72
Experiment 2											
IFR	.00	.01	.03	.11	.26	.38	.55	.63	.68	.79	.73
ISR-free	.00	.00	.03	.15	.33	.56	.60	.76	.80	.85	.90
Experiment 3											
ISR-free	.00	.01	.02	.12	.23	.49	.70	.78	.88	.88	.92
Standard ISR	.00	.00	.00	.06	.10	.20	.29	.38	.38	.48	.48
Experiment 4											
Free RoO	.00	.00	.05	.11	.28	.44	.51	.75	.74	.79	.87
Standard RoO	.00	.01	.02	.05	.12	.19	.21	.19	.12	.13	.14
E: Probability of first recall = "Other" serial positions											
Experiment 1											
IFR						.05	.10	.12	.11	.11	.17
Experiment 2											
IFR						.04	.10	.07	.13	.13	.22
ISR-free						.03	.07	.06	.05	.04	.04
Experiment 3											
ISR-free						.04	.05	.13	.05	.08	.07
Standard ISR						.13	.14	.17	.25	.30	.33
Experiment 4											
Free RoO						.00	.02	.02	.09	.09	.01
Standard RoO						.00	.06	.14	.29	.36	.41

Note. Table 6A summarizes the mean proportion of words recalled using free recall (FR) scoring in Experiments 1–3. Table 6B summarizes the mean proportion of words correctly assigned to the correct serial position using immediate serial recall (ISR) scoring in the relevant conditions of Experiments 2 to 4. Tables 6C, 6D, and 6E summarize the proportion of trials in which the first word recalled or selected was in (C) Serial Position 1, (D) one of the last four serial positions, or (E) one of the other serial positions. IFR = immediate free recall; RoO = reconstruction of order.

resultant serial position curves: When recall started at Serial Position 1, there was elevated recall of early list items; when recall started toward the end of the list, there were extended recency effects.

Toward Greater Theoretical Integration Between IFR and ISR

The similarities in these findings across the different tasks suggest that at least some of the differences observed between IFR and ISR result from the differences in the list lengths that are typically used in the two tasks. Thus, although IFR is typically associated with extended recency effects and recall starting with one of the last few items, such findings are specific to the longer list lengths that are typically used. Similarly, ISR is typically associated with extended primacy effects and recall starting with the first item; this finding is observed only at shorter list lengths.

However, our data across these different tasks suggest that there is a "natural" order in which to recall short lists of words, regardless of the exact task instructions: Participants will tend to start recall with the first list item, even when this is not strictly necessary, and continue to recall early list items. This finding provides firm experimental evidence consistent with earlier reports (Corballis, 1967; Neath & Crowder, 1996). As the list length increases, so participants show reduced ability to recall (or reconstruct) the first list item when instructed, resulting in reduced primacy and elevated recency effects.

These findings offer challenges to models of IFR (including our own) that assume that free recall is predominantly recency-based (e.g., Davelaar et al., 2005; Howard & Kahana, 2002a; Tan & Ward, 2000). The prevalence of recency in the IFR of longer lists (which are typically used) has perhaps overly focused theorizing on explaining recency effects, with primacy effects being relegated to control processes such as rehearsal or attention. Although one could argue that the early list items benefit from rehearsal or attention in this study, we suspect that we would obtain similar forward-ordered recall with shorter list lengths under revised methodology (such as with speeded presentation rates, articulatory suppression, or under divided attention, although these experiments have not as yet been performed).

We also believe that the findings offer challenges to models of ISR that are distinct from IFR (Baddeley, 1986, 2000; Baddeley & Hitch, 1974; Brown et al., 2000; Burgess & Hitch, 1992, 1999, 2006; Farrell & Lewandowsky, 2002; Henson, 1998; Lewandowsky & Farrell, 2008; Nairne, 1988; Oberauer & Lewandowsky, 2008; Page & Norris, 1998), especially given recent findings of similar effects of diagnostic variables such as presentation rate, word length, and articulatory suppression on the two tasks in lists of lengths 6 and 8 (Bhatarah et al., 2008, 2009). Rather, these findings suggest that there needs to be a closer relationship between theories of IFR and ISR.

Interpretations of Our Findings

One classic interpretation of our findings is that we are observing the encoding and output from STS in both IFR and ISR (Atkinson & Shiffrin, 1971; Waugh & Norman, 1965). Imagine an STS of limited capacity (say three to four items) that stores items in serial order from which the first item tends to be output first. At short list lengths, this mechanism will produce near perfect forward-ordered recall in all immediate memory tasks. As the list length increases, so the early list items held in this store would be increasingly overwritten by later list items, resulting in the decreased probability of starting recall with the first list item and the increased tendency to begin output with later list items, resulting in recency effects.

In support of an STS buffer account, we note that there are some similarities between the probabilities of first recall observed in our experiments and other classic immediate memory data. Recalling the first item in a list could be considered to be similar to a one-item Brown-Peterson task, where the first list item is the single word-to-be-remembered and the subsequent list items act as distractors. As discussed recently by Usher, Davelaar, Haarmann, and Goshen-Gottstein (2008), a buffer model would perhaps best predict a concave forgetting curve, which is what is shown in all of the PFR curves.

Although intuitively plausible, this STS account may potentially have difficulty in predicting the exact order of output from STS for both the primacy and the recency effects observed in the lists. To account for the primacy effect, the STS should be ordered so that the earliest items are output first. However, the recency item that is most commonly output first in free recall is the very last list item, potentially requiring different output preferences from STS when one of the last four recent items are output first. Moreover, a common STS account of recency and ISR should predict that the recency effect would be severely reduced if not eliminated by a concurrent six-digit ISR task, since both the recency items and the words stored for ISR should compete for the same limited-capacity STS. Contrary to this prediction, the magnitude of the recency effect in free recall is unaffected by a continuous or continual six-digit ISR task (Baddeley & Hitch, 1974, 1977; Bhatarah et al., 2006), at least when the free recall task is of a long list of 16 words. We therefore believe that there are difficulties for a common STS account of ISR and recency. However, if these potential problems could be overcome then better theoretical integration between IFR and ISR would be valuable in establishing a more universal and intuitive account of STS that supports performance in a wide range of immediate memory tasks.

A second interpretation of our findings is that the mechanisms for ISR and recency in IFR are rather different, such that there are distinct primacy-based mechanisms (typically associated with ISR) and separate recency-based mechanisms (typically associated with free recall), both of which operate at both tasks, but whose relative importance varies with increasing list length. According to this interpretation, the primacy-based mechanisms are increasingly dominant for shorter list lengths but are supplanted at longer list lengths by recency-based mechanisms. This account could explain the data by Baddeley and Hitch (1974, 1977; Bhatarah et al., 2006): ISR could be supported by an STS mechanism responsible for graded primacy, but the recency effect could reflect a different long-term memory system (and so not be affected catastrophically by the continual ISR task). Theoretical integration between IFR and ISR along these lines would remain valuable in helping to explain the IFR and ISR of List Lengths 4–9.

One potential problem for a mechanism that proposes graded primacy in IFR is that there is relatively little evidence of graded primacy for the first word recalled (e.g., Tables 2 and 3): There are relatively few trials in which the first word recalled is the second or the third word on the list, even for shorter lists. Rather, in IFR the first word recalled seems best described as either from Serial Position 1 or from one of the last four serial positions (and of these more commonly from more recent positions), suggesting that it is the first word in particular (rather than early words in general) that is especially accessible at the beginning of recall. This seems to be true for all of the immediate memory tasks with the exception of standard ISR, where there was a small but significant preference to initiate recall with Serial Position 2 rather than Serial Position 3.

A third interpretation of our findings is based on a unitary account of memory. It is clear that a recency-based account of free recall (e.g., Tan & Ward, 2000) cannot accommodate the data presented here without additional assumptions. This is because the tendency to initiate recall with the first list item at short list lengths (even in free recall) is simply too great. However, the data suggest that an account of free recall that could explain the patterns of initial recalls might also provide a ready-made account of ISR. One possibility that we are beginning to explore is that participants try to recall the first list item first on all tasks and for all list lengths.[3] If the accessibilities of the list items are recency-based, then the likelihood that the first item can be successfully retrieved will decrease with increasing list length, but for short lists, the chance of successfully recalling the first list item (indeed almost all items) may nevertheless be sufficiently high for recall to initiate with the first list item. If one additionally assumes that the recall of one item increases the probability of recalling its successor (e.g., Howard & Kahana, 2002a; Kahana, 1996), one might have the beginnings of a unitary account that could explain the data, albeit with an additional assumption. Clearly this is pure speculation at this stage, but this approach perhaps most accurately describes the first or graded recency function for the PFRs in IFR.

Outstanding Differences Between the Tasks

It is tempting in an article that aims to promote greater theoretical integration between IFR and ISR to emphasize the similarities in the data sets and downplay the differences. Despite many observed similarities between the different immediate memory tasks in these experiments, there remain a number of outstanding differences. First, participants were more likely to initiate recall with the first serial position at longer list lengths in tasks when they were required to do so (standard ISR and standard RoO) relative to tasks in which participants were free to output in any order (IFR, ISR-free, and free RoO). Second, at longer list lengths, the level of forward-ordered recall was greater in variants of ISR than in IFR. Finally, there was evidence for graded primacy in the PFR curves for standard ISR but not for any other task. An important question is whether these differences reflect the use of different memory mechanisms at encoding and/or retrieval in these tasks, or whether these differences are simply what one would expect if participants encoded the words using identical memory mechanisms but at retrieval were able to alter their output orders in line with different task instructions.

It is clear that participants have some capacity to vary their output orders with different task instructions, and this is perhaps best illustrated by recent studies (e.g., Bhatarah et al., 2008; Lewandowsky et al., 2008; Tan & Ward, 2007) that have manipulated the test expectancy of different task instructions, such that the method of testing is either known in advance (precued task instructions) or is known only immediately prior to test (postcued task instructions). For example, Bhatarah et al. (2008) compared pre- and postcued performance on IFR and standard ISR tasks, Tan and Ward (2007) compared pre- and postcued performance on standard ISR and ISR-free tasks, and Lewandowsky et al. (2008) compared pre- and postcued performance on standard RoO and free RoO tasks. In all these studies (which used list lengths of 6 items or more), participants tended to start with Serial Position 1 (resulting in enhanced primacy and reduced recency) when cued to do so, but they more often initiated recall with one of the last few list items when the postcue allowed free ordering of output (resulting in enhanced recency and reduced primacy).

Thus, it seems likely that some of the differences observed across the different immediate memory tasks reflect differences between what participants would *choose* to do when they have free choice and what participants are *required* to do by the task instructions. For example, the task requirement in standard ISR (to initiate recall with Serial Position 1) might lead participants to try to initiate recall with Serial Position 1 at longer list lengths than they would choose to do so, if they were free (as in IFR) to initiate recall in any order. Similarly, if participants could not recall the first serial position, they might try to recall Serial Position 2 more often in standard ISR (as instructed) than if they were free to initiate recall in any order. In addition, the task requirement to position the words in the correct grid position in variants of ISR might encourage participants at longer list lengths to retrieve longer runs of ordered recalls than they would choose to do so were they free (as in IFR) to recall without worrying about the items' serial positions.

However, any outstanding differences between the different immediate memory tasks might alternatively indicate the use of different memory mechanisms on different tasks. In line with this position, Lewandowsky et al. (2008) examined whether temporal isolation effects (TIEs; the recall advantage for items that are separated by larger temporal intervals) could be observed in the free RoO and the standard RoO tasks. Previous research by these authors had indicated that TIEs were present in tasks where participants had some freedom in output order, such as IFR (Brown, Morin, & Lewandowsky, 2006), but not in tasks with prescribed output orders, such as standard ISR (Lewandowsky, Brown, Wright, & Nimmo, 2006). In line with their previous findings, Lewandowsky et al. (2008) observed TIEs with free RoO (which like IFR allows participants some freedom in output order) but not standard RoO (which like standard ISR requires a totally prescribed output order). Importantly, the difference in sensitivity to TIEs in the two tasks was found when the method of testing was known in advance (precued) and also when the method of testing was not known in advance but was instead signaled after study (postcued). Lewandowsky et al. (2008) interpreted these findings as showing that the list items were encoded simultaneously along temporal and ordinal dimensions but that the different tasks made differential use of these different dimensions at test. For tasks such as free recall, ISR-free, or free RoO in which participants can choose their output order, the temporal dimension may be used, resulting in TIEs; whereas for tasks requiring strict forward-

[3] We wish to thank Gordon Brown for useful discussions leading to this line of enquiry.

ordered recall such as standard ISR and standard RoO, there are little or no TIEs because the ordinal dimension is used.

These findings suggest that a final interpretation of our data is that the lists of words in our experiments were also encoded simultaneously in multiple memory mechanisms (or along multiple dimensions), which were differentially used depending on task instructions and/or participant choice at test. These differences could usefully be addressed in future work using the methods and analyses presented here, with pre- and postcued conditions.

The Relationship Between Rehearsal and Recall in Free Recall

Finally, our data also go some way to explain the relationship between free recall and rehearsal. We have for some time been interested in whether the processes underpinning rehearsal could also be the same as those underpinning recall (see also Laming, 2006, 2008). In a number of earlier studies, we have presented participants with lists of 16 or 20 words for free recall and examined the patterns of rehearsals during study by asking the participants to rehearse out loud whatever list items they were thinking about during the interstimulus intervals (the overt rehearsal methodology; Rundus & Atkinson, 1970). One potential problem is why recall at the end of the list appears to be dominated by recency, but the rehearsal order early during list presentation tends to be forward-ordered and to show extended primacy effects (Tan & Ward, 2000; Ward et al., 2003). Indeed, Laming (2010) raised exactly this question—why should the early words be preferentially rehearsed in this way, given a recency-based account of memory? The data from the current experiment partially explain the differences between the primacy-based output orders for the early patterns of rehearsal and later recency-based patterns of recall. The patterns of rehearsal early in the study can be considered to be analogous to the IFR of words from a short list: forward serial recall, starting with the first list item. By contrast, the patterns of recall at the end of the list are from a longer list, where recall tends to be dominated by recalling first a recently experienced list item. It would appear that the patterns of rehearsals observed early in study in free recall closely resemble the patterns of recall of lists of words of similar length, providing evidence supporting the relationship between recall and rehearsal. We note that this insight does not fully explain the forward-ordered nature of rehearsal, nor the tendency to initiate rehearsals with the first list item, but rather demonstrates that the order of rehearsal and recall are similar and underscores the importance of understanding why the first list item is so preferentially accessed.

Summary and Concluding Thoughts

In summary, we believe that our data necessitate greater theoretical integration between IFR and ISR. Previous differences between IFR and ISR are at least partly due to the list lengths that are typically used. Most importantly, participants tend to initiate recall with the very first list item in all immediate memory tasks with short lists, even when they are not strictly required to do so. Moreover, the initial word recalled affects the subsequent patterns of output: ISR-like output occurs on both tasks when recall starts with Serial Position 1 (which is more likely with short lists), but free recall-like output occurs on both tasks when recall starts with one of the last four words (more likely with longer lists). Our data therefore strongly suggest that IFR and ISR are more similar than has been previously assumed and suggest that the focus on explaining primacy in ISR and recency in IFR needs to be readdressed: Both tasks can produce both serial position functions at different list lengths. Finally, we believe that our findings highlight the need for greater understanding as to how participants are able to gain preferential access to the first list item in immediate memory tasks.

References

Anderson, J. R., Bothell, D., Lebiere, C., & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38,* 341–380. doi:10.1006/jmla.1997.2553

Atkinson, R. C., & Shiffrin, R. M. (1971, August 1). The control of short-term memory. *Scientific American, 225,* 82–90. doi:10.1038/scientificamerican0871-82

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

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

Baddeley, A. (2000). The episodic buffer: A new component of working memory*? Trends in Cognitive Sciences, 4,* 417–423. doi:10.1016/S1364-6613(00)01538-2

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

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

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 & Cognition, 37,* 689–713. doi:10.3758/MC.37.5.689

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. doi:10.1037/0278-7393.32.2.215

Bhatarah, P., Ward, G., & Tan, L. (2008). Examining the relationship between free recall and immediate serial recall: The serial nature of recall and the effect of test expectancy. *Memory & Cognition, 36,* 20–34. doi:10.3758/MC.36.1.20

Botvinick, M. M., & Plaut, D. C. (2006). Short-term memory for serial order: A recurrent neural network model. *Psychological Review, 113,* 201–233. doi:10.1037/0033-295X.113.2.201

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

Brown, G. D. A., Chater, N., & Neath, I. (2008). Serial and free recall: Common effects and common mechanisms? A reply to Murdock (2008). *Psychological Review, 115,* 781–785. doi:10.1037/a0012563

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

Brown, G. D. A., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review, 114,* 539–576. doi:10.1037/0033-295X.114.3.539

Brown, G. D. A., Preece, T., & Hulme, C. (2000). Oscillator-based memory for serial order. *Psychological Review, 107,* 127–181. doi:10.1037/0033-295X.107.1.127

Burgess, N., & Hitch, G. (1992). Toward a network model of the articulatory loop. *Journal of Memory and Language, 31,* 429–460. doi:10.1016/0749-596X(92)90022-P

Burgess, N., & Hitch, G. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review, 106,* 551–581. doi:10.1037/0033-295X.106.3.551

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55,* 627–652. doi:10.1016/j.jml.2006.08.005

Corballis, M. C. (1967). Serial order in recognition and recall. *Journal of Experimental Psychology, 74,* 99–105. doi:10.1037/h0024500

Crannell, C. W., & Parrish, J. N. (1957). A comparison of immediate memory span for digits, letters, and words. *Journal of Psychology, 44,* 319–327.

Crowder, R. G. (1969). Behavioral strategies in immediate memory. *Journal of Verbal Learning and Verbal Behavior, 8,* 524–528. doi:10.1016/S0022-5371(69)80098-8

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. doi:10.1037/0033-295X.112.1.3

Drewnowski, A., & Murdock, B. B., Jr. (1980). The role of auditory features in memory span for words. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 319–332. doi:10.1037/0278-7393.6.3.319

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

Farrell, S., & Lewandowsky, S. (2008). Empirical and theoretical limits on lag-recency in free recall. *Psychonomic Bulletin & Review, 15,* 1236–1250. doi:10.3758/PBR.15.6.1236

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.

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, NY: Academic Press.

Golomb, J. D., Peelle, J. E., Addis, K. M., Kahana, M. J., & Wingfield, A. (2008). Effects of adult aging on utilization of temporal and semantic associations during free and serial recall. *Memory & Cognition, 36,* 947–956. doi:10.3758/MC.36.5.947

Grossberg, S., & Pearson, L. R. (2008). Laminar cortical dynamics of cognitive and motor working memory, sequence learning and performance: Toward a unified theory of how the cerebral cortex works. *Psychological Review, 115,* 677–732. doi:10.1037/a0012618

Healy, A. F. (1982). Short-term memory for order information. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 191–238). New York, NY: Academic Press.

Healy, A. F., Fendrich, D. W., Cunningham, T. F., & Till, R. E. (1987). Effects of cuing on short-term retention of order information. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 413–425. doi:10.1037/0278-7393.13.3.413

Henson, R. N. A. (1998). Short-term memory for serial order: The start–end model of serial recall. *Cognitive Psychology, 36,* 73–137. doi:10.1006/cogp.1998.0685

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. doi:10.1037/0278-7393.25.4.923

Howard, M. W., & Kahana, M. J. (2002a). A distributed representation of temporal context. *Journal of Mathematical Psychology, 46,* 269–299. doi:10.1006/jmps.2001.1388

Howard, M. W., & Kahana, M. J. (2002b). When does semantic similarity help episodic retrieval? *Journal of Memory and Language, 46,* 85–98. doi:10.1006/jmla.2001.2798

Howard, M. W., Sederberg, P. B., & Kahana, M. J. (2009). Reply to Farrell & Lewandowsky: Recency-contiguity interactions predicted by the temporal context model. *Psychonomic Bulletin & Review, 16,* 973–984. doi:10.3758/PBR.16.5.973

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

Kahana, M. J., Howard, M. W., & Polyn, S. M. (2008). Associative retrieval processes in episodic memory. In J. Byrne (Series Ed.), *Learning and memory: A comprehensive reference. Vol. 2: Cognitive psychology of memory* (H. L. Roediger, III, Vol. Ed.). Oxford, England: Elsevier. doi:10.1016/B978-012370509-9.00185-6

Klein, K. A., Addis, K. M., & Kahana, M. J. (2005). A comparative analysis of serial and free recall. *Memory & Cognition, 33,* 833–839.

Laming, D. (1999). Testing the idea of distinct storage mechanisms in memory. *International Journal of Psychology, 34,* 419–426. doi:10.1080/002075999399774

Laming, D. (2006). Predicting free recalls. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1146–1163. doi:10.1037/0278-7393.32.5.1146

Laming, D. (2008). An improved algorithm for predicting free recalls. *Cognitive Psychology, 57,* 179–219. doi:10.1016/j.cogpsych.2008.01.001

Laming, D. (2009). Failure to recall. *Psychological Review, 116,* 157–186. doi:10.1037/a0014150

Laming, D. (2010). Serial position curves in free recall. *Psychological Review, 117,* 93–133. doi:10.1037/a0017839

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

Lewandowsky, S., Brown, G. D. A., Wright, T., & Nimmo, L. M. (2006). Timeless memory: Evidence against temporal distinctiveness models of short term memory for serial order. *Journal of Memory and Language, 54,* 20–38. doi:10.1016/j.jml.2005.08.004

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. *The Psychology of Learning and Motivation, 49,* 1–48. doi:10.1016/S0079-7421(08)00001-7

Lewandowsky, S., Nimmo, S., & Brown, G. D. A. (2008). When temporal isolation benefits memory for serial order. *Journal of Memory and Language, 58,* 415–428. doi:10.1016/j.jml.2006.11.003

Mandler, G., & Dean, P. J. (1969). Seriation: The development of serial order in free recall. *Journal of Experimental Psychology, 81,* 207–215. doi:10.1037/h0027767

Mandler, G., Worden, P. E., & Graesser, A. C. (1974). Subjective disorganization: Search for the locus of list organization. *Journal of Verbal Learning and Verbal Behavior, 13,* 220–235. doi:10.1016/S0022-5371(74)80047-2

Miller, G. A. (1956). The magic number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97. doi:10.1037/h0043158

Murdock, B. B., Jr. (1962). The serial position effect of free recall. *Journal of Experimental Psychology, 64,* 482–488. doi:10.1037/h0045106

Murdock, B. (2008). Issues with the SIMPLE model: Comment on Brown, Neath, and Chater (2007). *Psychological Review, 115,* 779–780. doi:10.1037/0033-295X.115.3.779

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. doi:10.1111/j.1467-9280.1992.tb00027.x

Nairne, J. S., & Neumann, C. (1993). Enhancing effects of similarity on long-term memory for order. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 329–337. doi:10.1037/0278-7393.19.2.329

Neath, I. (1997). Modality, concreteness, and set-size effects in a free reconstruction of order task. *Memory & Cognition, 25,* 256–263.

Neath, I., & Crowder, R. G. (1996). Distinctiveness and very short-term serial position effects. *Memory, 4,* 225–242. doi:10.1080/096582196388933

Oberauer, K., & Lewandowsky, S. (2008). Forgetting in immediate serial recall: Decay, temporal distinctiveness, or interference? *Psychological Review, 115,* 544–576. doi:10.1037/0033-295X.115.3.544

Page, M. P. A., & Norris, D. (1998). The primacy model: A new model of immediate serial recall. *Psychological Review, 105,* 761–781. doi:10.1037/0033-295X.105.4.761-781

Pellegrino, J. W., & Battig, W. F. (1974). Relationships among higher order organizational measures and free recall. *Journal of Experimental Psychology, 102,* 463–472. doi:10.1037/h0035898

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. doi:10.1037/a0014420

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review, 88,* 93–134. doi:10.1037/0033-295X.88.2.93

Rundus, D. (1971). Analysis of rehearsal processes in free recall. *Journal of Experimental Psychology, 89,* 63–77. doi:10.1037/h0031185

Rundus, D., & Atkinson, R. C. (1970). Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning and Verbal Behavior, 9,* 99–105. doi:10.1016/S0022-5371(70)80015-9

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. doi:10.1037/a0013396

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. doi:10.1037/0278-7393.26.6.1589

Tan, L., & Ward, G. (2007). Output order in immediate serial recall. *Memory & Cognition, 35,* 1093–1106.

Tulving, E. (1962). Subjective organization in free recall of "unrelated" items. *Psychological Review, 69,* 344–354. doi:10.1037/h0043150

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. doi:10.1037/a0013725

Wallace, W. P. (1970). Consistency of emission order in free recall. *Journal of Verbal Learning and Verbal Behavior, 9,* 58–68. doi:10.1016/S0022-5371(70)80008-1

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

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. doi:10.1037/0278-7393.29.2.186

Waugh, N. C., & Norman, D. A. (1965). Primary memory. *Psychological Review, 72,* 89–104. doi:10.1037/h0021797

Waugh, N. C. (1960). Serial position and the memory span. *American Journal of Psychology, 73,* 68–78. doi:10.2307/1419117

Received November 30, 2009
Revision received March 22, 2010
Accepted April 12, 2010 ■