Beginning at the beginning:

Recall order and the number of words to be recalled

Lydia Tan[1], Geoff Ward[2], Laura Paulauskaite[1] and Maria Markou[1]

[1]City University London

[2]University of Essex

Correspondence to:
Lydia Tan,
Department of Psychology,
City University London,
Northampton Square,
London EC1V 0HB,
U.K.

Telephone: +44 207 0408379
E-mail: Lydia.Tan.1@city.ac.uk

Abstract

When participants are asked to recall a short list of words in any order that they like, they tend to initiate recall with the first list item and proceed in forwards order, even when this is not a task requirement. The current research examined whether this tendency might be influenced by varying the number of items that are to be recalled. In three experiments, participants were presented with short lists of between 4 and 6 words and instructed to recall 1, 2, 3 or all of the items from the lists. Data were collected using immediate free recall (IFR, Experiment 1), immediate serial recall (ISR, Experiment 2) and a variant of ISR that we call ISR-free (Experiment 3), in which participants had to recall words in their correct serial positions but were free to output the words in any order. For all three tasks, the tendency to begin recall with the first list item occurred only when participants were required to recall as many items from the list as they could. When participants were asked to recall only one or two items, they tended to initiate recall with end-of-list items. It is argued that these findings show for the first time a manipulation that eliminates the initial tendency to recall in forward order, provide some support for recency-based accounts of IFR and help explain differences between single-response and multiple-response immediate memory tasks.

(232 words)

In recent years, there has been a growing belief that much could be gained by the theoretical integration of two widely used and highly influential immediate memory tasks, immediate serial recall (ISR) and immediate free recall (IFR) (e.g., Anderson, Bothell, Lebiere, & Matessa, 1998; Bhatarah, Ward & Tan, 2006, 2008; Brown, Chater & Neath, 2008; Brown, Neath & Chater, 2007; Farrell, 2012; Grenfell-Essam & Ward, 2012; Grossberg & Pearson, 2008; Hurlstone, Hitch & Baddeley, 2014; Kahana, 2012; Klein, Addis & Kahana, 2005; Ward, Tan & Grenfell-Essam, 2010). In these tasks, participants are presented with lists of words and at the end of each list, they are either free to recall as many of the list items as possible in any order that they like (IFR) or they are required to recall as many items as possible in exactly the same serial order as that in which they were presented (ISR).

One reason to believe that such integration is possible is the observation that participants tend to output their recalls in forwards serial order, even in IFR tasks (e.g., Beaman & Morton, 2000; Bhatarah et al., 2008; Golomb, Peelle, Addis, Kahana, & Wingfield, 2008; Howard & Kahana, 1999; Kahana, 1996; Klein et al., 2005; Laming, 1999, 2006, 2008, 2010; Ward et al., 2010), a finding that has led some to suggest that forward-ordered recall may be a defining property of episodic memory (e.g., Hurlstone et al., 2014).

A second reason for believing that such integration is possible is that when IFR and ISR are compared under identical methodological conditions, list lengths and scoring systems, both tasks are similarly affected by speech-based variables such as word length, presentation rate and concurrent articulation (Bhatarah, Ward, Smith & Hayes, 2009), as well as concurrent articulation and phonological similarity (Spurgeon, Ward & Matthews, 2014b). Moreover, both tasks are encoded (and rehearsed) in similar ways, with little or no significant effect on the serial position curves and output orders, irrespective of whether or not advance

knowledge has been given as to which of the two tasks is to be performed (e.g., Bhatarah et al., 2008; 2009; Grenfell-Essam & Ward, 2012).

Finally, when participants are presented with lists of between one and 15 words in a range of memory tasks, including IFR and ISR, participants tend to initiate recall of short lists with the first list item, but of longer lists with one of the last few items (e.g., Cortis, Dent, Kennett & Ward, 2015; Grenfell-Essam & Ward, 2012; Spurgeon, Ward & Matthews, 2014a; Ward et al., 2010). For example, when Ward et al. (2010) presented participants with a short list of words such as "victim, hollow, future, kitten" to recall *in any order* (IFR), they tended to recall "victim, hollow, future, kitten"; that is, they recalled in an "ISR-like" manner, even though this was not necessitated by the task instructions (see also Corballis, 1967; Neath & Crowder, 1996). Ward et al. found that the position within the list at which one initiated recall had large effects on the resultant serial position curves: if recall started with the first list item, there was likely to be elevated recall of early list items and reduced recency, whereas if recall started with one of the last list items, there was likely to be elevated recall of later list items and reduced primacy.

Ward et al. not only interpreted these findings as evidence that ISR and IFR may be more similar than had been previously assumed, but they also proposed that such findings presented difficulties for many unitary theories of IFR that have assumed that the serial position curves and the probability of first recall (PFR) data in IFR would be dominated by extended recency effects. Although the serial position curves and the PFR data correctly support the supremacy of recency effects over primacy effects with longer lists, temporal distinctiveness accounts of IFR propose that the most recent list items will always be those that are most temporally distinctive (e.g., Brown et al., 2007; Glenberg & Swanson, 1986). Similarly, accounts that assume that items are associated with a drifting or changing temporal context (e.g., Howard & Kahana, 2002; Polyn, Norman & Kahana, 2009; Sederberg, Howard

& Kahana, 2008; Tan & Ward, 2000) propose that more recent items should be generally more readily accessible than earlier list items, because they share contexts that are more closely related to the test context. The finding that it is the first item (and not the last item) that is most accessible when recalling short lists creates a distinct deviation from the predictions of these unitary accounts. The generality of this finding and the constraints it places on theories of episodic memory suggest that this phenomenon may be a benchmark finding that such accounts should be able to explain.

Why, then, do participants initiate recall of short lists with the first item? Our attempts to date at answering this question have mostly allowed us only to rule out some of the possible reasons. The finding is remarkably robust and can be obtained under concurrent articulatory suppression and at fast presentation rates (Grenfell-Essam, Ward & Tan, 2013), suggesting that it is not caused by rehearsal. The finding is present, although somewhat attenuated, with free recall under continual distractor conditions and delayed free recall conditions (Spurgeon et al., 2014a), suggesting that it is not due to the immediate output of a short-term buffer store. It is also present but similarly attenuated with visual presentation under concurrent articulatory suppression (Spurgeon et al., 2014b), suggesting that it is not due to the proposed function of the phonological store or phonological loop (Baddeley, 1986). It is even present with the immediate recall of visual-spatial dots and tactile stimulations to the face (Cortis et al., 2015), suggesting that the finding is not the result of an exclusively verbal mechanism. The finding is also present in lists preceded by a stimulus prefix and observed in methods requiring the allocation of attention to other list items (Grenfell-Essam & Ward, 2015). It is also observed when the first item in a short list is not within the current temporal group (Spurgeon, Ward & Farrell, 2015), suggesting that temporal grouping does not provide a complete account. We note that although the tendency to initiate recall with the first list item has been shown to be somewhat attenuated under some

of these conditions, it nevertheless always remains the modal tendency at the shortest list lengths.

Given the apparent robustness and ubiquity of this finding, we believe an important insight might be revealed if a manipulation could be found that would eliminate this tendency. One such manipulation is suggested by the patterns of recall that are present in certain other immediate memory tasks, such as recognition in a modified Sternberg task (Oberauer, 2003), old-new recognition (as measured by reaction times; Duncan & Murdock, 2000) and some probed recall studies (e.g., Avons, Wright & Pammer, 1994; Penney, 1982). The serial position curves obtained in these paradigms are rather different from those typically observed with the ISR and IFR of short lists because they are almost entirely dominated by recency effects, with few if any primacy effects.

One obvious difference between the methodologies of these two sets of tasks is that ISR and IFR require participants to recall all the list items, whereas yes-no recognition and probed recall require participants to recall only a single list item. When participants are presented with a short list of words and are post-cued to perform either ISR or single yes-no recognition (Duncan & Murdock, 2000), the serial position curves in the post-cued yes-no recognition task differ markedly from those obtained when the task is pre-cued, that is, predictably known in advance. This suggests that requiring participants to recall all the items may lead to a different encoding strategy from requiring them to make a single response. Moreover, it has been demonstrated that when serial recall and serial yes-no recognition tasks are both performed with multiple responses and the need for positional information, the serial position curves are very similar for these tasks and show large primacy and small recency effects (Oberauer, 2003).

Could it be, therefore, that one reason why participants have tended to initiate recall of short lists with the first item in the various tasks employed is that, to date, all the

methodologies associated with such a finding have required participants to recall as many of the list items as possible (e.g., Grenfell-Essam & Ward, 2012; Ward et al., 2010)? If participants are given a short list of items, such as "victim, hollow, future, kitten", would they still begin their recalls with "victim" when asked to recall just one of the list items? If not "victim", which item are they most likely to recall instead? Given the above line of reasoning, it would not be unreasonable to suppose that single-item responses might be dominated by recency effects, such that participants might very well initiate their recalls with the last item when asked to recall only one of the list items.

The present series of experiments therefore aimed to determine whether the tendency to spontaneously initiate recall of short lists of words with the first list item occurs only when participants are required to recall all the list items. By manipulating the number of items that are to be recalled (one, two, three, or all) and the number of words in the list (four, five, or six), we would be able to examine whether there are any discernible patterns in participants' order of recall across the number of items to be recalled and list length. We would expect to replicate the Ward et al. (2010) finding that participants tend to initiate their recalls with the first list item when they are asked to recall all the items and we would expect this tendency to decrease with increasing list length. What is unknown is whether this tendency is greatly reduced or even eliminated when participants are asked to recall fewer list items. Under these conditions, our consideration of the yes-no recognition and probed recall tasks suggests we may see an increasing tendency for participants to preferentially access recency items.

Experiment 1

The free recall paradigm was used in Experiment 1. In each trial, participants were shown a cue specifying the number of words they were to recall for the upcoming list (one, two, three, or all), followed by 4, 5 or 6 words for IFR. List length was randomized so that

participants did not know the length of each list in advance of the recall cue and it was difficult to predict when a given list was going to end. Consequently, although the number of words to be recalled was always pre-cued, participants would be unlikely to use an encoding strategy specific to list length.

At the end of the list, participants had to recall the requisite number of words by writing them down on a response grid. Of interest was whether participants would be more likely to begin recalling from the end of the list when asked for fewer items to recall.

Method

Participants. Twenty-four psychology students from City University London participated in this experiment in exchange for course credits. All participants were able to read and write English fluently and had normal or corrected-to-normal vision.

Materials and apparatus. The materials consisted of 600 monosyllabic words randomly selected from the MRC Psycholinguistic Database (Coltheart, 1981), with frequencies of occurrence of 10 per million and above, based on the Kučera and Francis (1967) norms. From these 600 words, 120 experimental lists were constructed, 40 for each of the list lengths of 4, 5 or 6 words. The composition of each list was randomized for each participant and no participant saw the same word twice. A response sheet with 120 boxes, each having six numbered lines, was provided to the participants for free recall. The words were presented in 24-point Courier New bold font on a computer monitor using the E-prime application.

Design. A within-subjects design was used. There were three within-subjects independent variables: recall requirement with 4 levels (recall 1, recall 2, recall 3, recall all), list length with 3 levels (4, 5, 6), and serial position (SP) with up to 6 levels (SPs 1-6). The

main dependent variable was the probability that an item from a given serial position was recalled first (the probability of first recall (PFR)).

   Procedure. Participants were tested individually and were presented with two practice trials, the first of five words and the second of four words, followed by 120 experimental word lists. The experimental trials were arranged into four blocks, one for each of the four recall requirement conditions. The order of the blocks was randomized across participants. Each block contained 30 trials, consisting of 10 trials each of the three different list lengths. The list lengths within each block were randomized. Each trial began with a prompt which informed the participants of the number of words from the upcoming list they should recall. After two seconds, a series of 4, 5 or 6 words was presented one at a time in the centre of the computer screen and participants read each word aloud as it was presented. Each word was displayed for two seconds. At the end of each list, a visual cue appeared, instructing the participants to recall either all the words or only one, two, or three words from the list they had just seen, in any order that they wished. Participants wrote down their responses on a response sheet. Recall was self-paced and participants were given as much time as they required to complete their recalls.

Results

   The probabilities of first recall (PFRs) for each list length, recall requirement and serial position are presented in Figure 1. (The PFR refers to the proportion of trials in which the first word recalled was from a particular serial position.)

--Figure 1 about here--

An inspection of Figures 1A, 1B and 1C suggests that for all three list lengths, participants were most likely to initiate their recall with the first list item when asked to recall all the items in the list. However, this tendency decreased slightly as list length increased. The tendency to initiate recall with the final list item was greatest when participants were asked to recall only one list item; this tendency remained relatively constant across the three list lengths.

The PFR data were analyzed by performing separate 3 (list length: 4, 5, 6) x 4 (recall requirement: recall 1, recall 2, recall 3 and recall all) within-subjects ANOVAs for the first, final, penultimate and antepenultimate serial positions. Figure 2 shows this PFR data for each list length and recall condition.

In all analyses throughout this paper, the degrees of freedom were corrected using the Greenhouse-Geisser correction wherever the assumption of sphericity was violated.

--Figure 2 about here--

First serial position. There was a significant effect of list length, $F(1.48, 33.96) = 53.85$, $p < .001$, $\eta_p^2 = .701$, a significant main effect of recall requirement, $F(3, 69) = 15.08$, $p < .001$, $\eta_p^2 = .396$, and a significant interaction effect between list length and recall requirement, $F(4.03, 92.72) = 5.90$, $p < .001$, $\eta_p^2 = .204$. Simple main effects revealed that for list length 4, the "recall 3" and "recall all" conditions were significantly different from each other and from the other recall conditions (all ps at least $< .05$). For list lengths 5 and 6, the "recall all" condition was significantly different from all other recall conditions (all ps at least $< .05$). Simple main effects also revealed that for the "recall 2" condition, list lengths 4 and 6 were significantly different from each other ($p < .01$); Figure 2A shows that the PFR was greater for list length 4 than for list length 6. Finally, for the "recall 3" and "recall all"

conditions, all three list lengths were significantly different from one another (all ps at least $<$ .05). Figure 2A indicates that for these recall conditions, the PFRs decreased as list length increased.

Final serial position. There was a non-significant effect of list length, $F(2, 46) = 1.81$, $p > .05$, $\eta_p^2 = .073$, a significant main effect of recall requirement, $F(1.80, 41.48) = 6.42$, $p < .01$, $\eta_p^2 = .218$, and a significant interaction effect between list length and recall requirement, $F(6, 138) = 2.54$, $p < .05$, $\eta_p^2 = .099$. Simple main effects revealed that for list length 4, the "recall 1" condition was significantly different from all other recall conditions (all ps at least $< .05$). For list length 6, the "recall 1" and "recall 2" conditions were significantly different from each other ($p < .01$).

Penultimate serial position. There was a non-significant effect of list length, $F(2, 46)$ $= .10$, $p > .05$, $\eta_p^2 = .004$, a significant main effect of recall requirement, $F(1.60, 36.72) = 7.17$, $p < .01$, $\eta_p^2 = .238$, and a significant interaction effect between list length and recall requirement, $F(6, 138) = 2.83$, $p < .05$, $\eta_p^2 = .109$. Simple main effects revealed that for list length 4, the "recall all" condition was significantly different from the "recall 1" and "recall 2" conditions (ps at least $< .05$), and the "recall 2" and "recall 3" conditions were significantly different from each other ($p < .01$). Finally, for list length 5, the "recall 2" condition was significantly different from the "recall all" condition ($p < .01$).

Antepenultimate serial position. There was a non-significant effect of list length, $F(2, 46) = 2.81$, $p > .05$, $\eta_p^2 = .109$, a non-significant main effect of recall requirement, $F(2.04, 46.90) = 2.89$, $p > .05$, $\eta_p^2 = .112$, and a non-significant interaction effect between list length and recall requirement, $F(6, 138) = .83$, $p > .05$, $\eta_p^2 = .035$.

Discussion

The results from Experiment 1 suggest that in IFR, a major factor which determines the position in the list from which participants initiate their recall is the number of items they are asked to recall. Consistent with Ward et al, (2010), in all three list lengths, the tendency to initiate recall with the first list item was highest when participants were required to recall all the words from the list and this effect appears to decrease with increasing list length. However, this tendency decreased when participants were required to recall only 1, 2 or 3 items. Indeed, the modal tendency was to initiate recall with the last list item when participants were asked to recall a single item, and the penultimate item when participants were asked to recall two items. This initiation of recall with the last or penultimate list items when required to recall only 1 or 2 items, respectively, must reflect a retrieval strategy, as it is present at all list lengths, even when the end of the list was not known at the time of encoding. This suggests that participants have privileged access to the first, last, and in some analyses penultimate list items, but they do not appear to have unlimited flexibility at retrieval, as there was little tendency to initiate recall with the antepenultimate item when participants were asked to recall 3 list items.

Experiment 2

The recall requirement and list length manipulations of Experiment 1 were repeated in Experiment 2, using the ISR task. Given that participants in both ISR and IFR tend to initiate recall with the first list item for short lists and one of the last four items for long lists (Ward et al., 2010), we were interested to discover whether ISR was similarly affected by the recall demands of the task, as was demonstrated in Experiment 1.

Method

Participants. Twenty-four psychology students from City University London participated in this experiment in exchange for course credits. All participants were able to read and write English fluently and had normal or corrected-to-normal vision. None had taken part in Experiment 1.
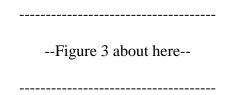
Materials and apparatus. The materials and apparatus were identical to those used in Experiment 1.

Design. There were three within-subjects independent variables: recall requirement, with 4 levels (recall 1, recall 2, recall 3, recall all); list length, with 3 levels (4, 5, 6); and serial position, with up to 6 levels (1-6). The main dependent variable was the PFR for each serial position.

Procedure. The procedure was identical to that used in Experiment 1, with the exception that participants carried out ISR instead of free recall at the end of each list. Following the end of the list, they were required to write down their responses in strict forward serial order, working down the response grid and writing each word on a row that corresponded to its serial position at presentation. Participants were told to leave a blank for any words they could not recall.

Results

The probabilities of first recall (PFRs) for each list length, recall requirement and serial position are presented in Figure 3.

--Figure 3 about here--

Figures 3A, 3B and 3C show broadly similar patterns across the three list length conditions: Participants were most likely to begin their recall with the first list item when

required to recall all the list items. In contrast, they were most likely to initiate their recall with the final list item when asked to recall only one item from the list. In addition, there was a pronounced tendency to initiate recall with the penultimate list item when asked to recall two items from the list.

As in Experiment 1, the PFR data were analyzed by performing separate 3 (list length: 4, 5, 6) x 4 (recall requirement: recall 1, recall 2, recall 3 and recall all) within-subjects ANOVAs for the first, final, penultimate and antepenultimate serial positions. Figure 4 shows this PFR data for each list length and recall condition.

--Figure 4 about here--

First serial position. There was a significant effect of list length, $F(2, 46) = 35.43$, $p < .001$, $\eta_p^2 = .606$, a significant main effect of recall requirement, $F(3, 69) = 60.45$, $p < .001$, $\eta_p^2 = .724$, and a significant interaction effect between list length and recall requirement, $F(6, 138) = 3.26$, $p < .01$, $\eta_p^2 = .124$. Simple main effects revealed that for list lengths 4 and 5, the "recall 3" and "recall all" conditions were significantly different from each other and from the other recall conditions (all ps at least $< .01$). For list length 6, the "recall all" condition was significantly different from all other recall conditions (ps $< .001$). Simple main effects also revealed that for the "recall 2" condition, list lengths 4 and 6 were significantly different from each other ($p < .05$); Figure 4A indicates that the PFR for list length 4 was greater than that for list length 6. For the "recall 3" condition, list length 4 was significantly different from list lengths 5 and 6 (ps at least $< .01$). Figure 4A reveals that the PFR was greater for list length 4 than for the other two list lengths. Finally, for the "recall all" condition, all three list lengths were significantly different from one another (all ps at least $< .05$). Figure 4A demonstrates that the PFR decreased as list length increased.

Final serial position. There was a non-significant effect of list length, $F(2, 46) = .02$, $p > .05$, $\eta_p^2 = .001$, a significant main effect of recall requirement, $F(1.01, 23.14) = 68.18$, $p < .001$, $\eta_p^2 = .748$, and a non-significant interaction effect between list length and recall requirement, $F(6, 138) = .04$, $p > .05$, $\eta_p^2 = .002$. Post-hoc pairwise comparisons revealed that the "recall 1" condition was significantly different from all other recall conditions ($ps < .001$).

Penultimate serial position. There was a non-significant effect of list length, $F(2, 46) = 1.70$, $p > .05$, $\eta_p^2 = .069$, a significant main effect of recall requirement, $F(1.42, 32.55) = 55.11$, $p < .001$, $\eta_p^2 = .706$ and a non-significant interaction effect between list length and recall requirement, $F(3.21, 73.88) = 1.46$, $p > .05$, $\eta_p^2 = .060$. Post-hoc pairwise comparisons revealed that the "recall 2" condition was significantly different from all other recall conditions ($ps < .001$).

Antepenultimate serial position. There was a non-significant effect of list length, $F(2, 46) = .25$, $p > .05$, $\eta_p^2 = .011$, a significant main effect of recall requirement, $F(1.82, 41.79) = 18.97$, $p < .001$, $\eta_p^2 = .452$, and a significant interaction effect between list length and recall requirement, $F(3.35, 77.02) = 3.22$, $p < .05$, $\eta_p^2 = .123$. Simple main effects revealed that for list length 4, the "recall 3" condition was significantly different from the "recall 2" and "recall all" conditions (ps at least $< .05$). For list lengths 5 and 6, the "recall 3" condition was significantly different from all other recall conditions (all ps at least $< .05$).

Discussion

The results from Experiment 2 reveal similar patterns across the three list length conditions: whichever item participants choose to recall first is greatly influenced by the number of items they are required to recall. Participants are most likely to start with the first list item when asked to recall all the items and are most likely to start with the last list item when asked to recall only one item. For all three list lengths, the likelihood with which the

first list item is recalled first (i.e., the PFR for serial position 1) decreases with the number of items to be recalled. In addition, the tendency to begin recalling with the first item generally decreases as list length increases, a finding consistent with previous research (e.g., Ward et al., 2010).

While these findings are noteworthy and serve to demonstrate in conjunction with Experiment 1 that the patterns of recall in ISR and IFR are similar, our conclusions concerning the "recall all" condition must be tempered by the consideration that in the ISR task, participants must necessarily begin recalling from the first list item if they are to recall all of the items as instructed. It is of interest to examine how participants would recall when their recall order is unconstrained, even while they are required to maintain positional information about the list items.

Experiment 3

The methodology used in Experiment 3 uses the "ISR-free" task (see Ward et al., 2010). This task has been used previously by us under the name "ISR with free output order" (Tan & Ward, 2007) and it has historically been used in prior research (e.g., Crowder, 1969; Waugh, 1960). In this task, participants are presented with a list of words and are then required to recall the words in their correct serial positions. Positional information is therefore required, but unlike in standard ISR, participants in the ISR-free task are free to recall the items in any (temporal) order they wish. The advantage of using such a methodology is that it enables us to determine which items participants would choose to recall first while maintaining the requirement, as in ISR, that they recall the input serial position of the presented items.

Method

Participants. As in the previous two experiments, twenty-four psychology students from City University London participated in this experiment in exchange for course credits. All participants were able to read and write English fluently and had normal or corrected-to-normal vision. None had taken part in the previous experiments.
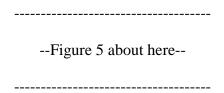
Materials and apparatus. The materials and apparatus were identical to those used in Experiments 1 and 2.

Design. There were three within-subjects independent variables: recall requirement, with 4 levels (recall 1, recall 2, recall 3 and recall all), list length, with 3 levels (4, 5, 6), and serial position, with up to 6 levels (1-6). The main dependent variable was the PFR for each serial position.

Procedure. The procedure was identical to that used in Experiments 1 and 2, with the exception that participants performed the ISR-free task instead of IFR at the end of each list. In this method, participants were free to write down their responses on the response grid in any temporal order they wished, but had to ensure that each word was written on a row that corresponded to its serial position at presentation. Participants were told to leave a blank for any words they could not recall.

Results

The probabilities of first recall (PFRs) for each list length, recall requirement and serial position are presented in Figure 5.

--Figure 5 about here--

Figures 5A, 5B and 5C suggest that, once again, participants are most likely to initiate their recall with the first list item in the "recall all" condition and to initiate their recall with

the final list item in the "recall one" condition. There is also a marked tendency to initiate recall with the penultimate item in the "recall two" condition.

As in the previous two experiments, the PFR data were analyzed by performing separate 3 (list length: 4, 5, 6) x 4 (recall requirement: recall 1, recall 2, recall 3 and recall all) within-subjects ANOVAs for the first, final, penultimate and antepenultimate serial positions. Figure 6 shows this PFR data for each list length and recall condition.

--Figure 6 about here--

First serial position. There was a significant effect of list length, $F(2, 46) = 67.42$, $p < .001$, $\eta_p^2 = .746$, a significant main effect of recall requirement, $F(3, 69) = 25.01$, $p < .001$, $\eta_p^2 = .521$, and a significant interaction effect between list length and recall requirement, $F(6, 138) = 10.94$, $p < .001$, $\eta_p^2 = .322$. Simple main effects revealed that for list length 4, all recall conditions were different from all other recall conditions (all ps at least $< .05$). For list length 5, the "recall 1" condition was significantly different from all other recall conditions (all ps at least $< .05$), and the "recall 2" condition was significantly different from the "recall all" condition ($p < .05$). Simple main effects also revealed that for the "recall 2", "recall 3" and "recall all" conditions, all three list lengths were significantly different from one another (all ps at least $< .05$). An inspection of Figures 6A reveals that the PFRs for these recall conditions decreased as list length increased.

Final serial position. There was a significant effect of list length, $F(2, 46) = 4.22$, $p < .05$, $\eta_p^2 = .155$, a significant main effect of recall requirement, $F(1.31, 30.22) = 46.95$, $p < .001$, $\eta_p^2 = .671$, and a significant interaction effect between list length and recall requirement, $F(3.78, 86.92) = 2.91$, $p < .05$, $\eta_p^2 = .112$. Simple main effects revealed that for list lengths 4, 5 and 6, the "recall 1" condition was significantly different from all other recall

conditions (all ps at least $< .01$). Simple main effects also revealed that for the "recall 2" condition, list lengths 4 and 6 were significantly different from each other ($p < .05$). Figure 6B shows that the PFR for list length 6 was greater than that for list length 4. Finally, for the "recall all" condition, list length 4 was significantly different from the other two list lengths (ps $< .01$). Figure 6B indicates that the PFR for list length 4 was smaller than for list lengths 5 and 6.

Penultimate serial position. There was a significant effect of list length, $F(2, 46) = 7.15$, $p < .01$, $\eta_p^2 = .237$, a significant main effect of recall requirement, $F(2.14, 49.12) = 32.45$, $p < .001$, $\eta_p^2 = .585$, and a non-significant interaction effect between list length and recall requirement, $F(6, 138) = 1.81$, $p > .05$, $\eta_p^2 = .073$. Post-hoc pairwise comparisons revealed that list length 6 was significantly different from the other two list lengths (ps $< .05$). Figure 6C demonstrates that the PFR was larger for list length 6 than for list lengths 4 and 5. In addition, the "recall 2" condition was significantly different from all other recall conditions (ps $< .001$).

Antepenultimate serial position. There was a significant effect of list length, $F(2, 46) = 7.32$, $p < .01$, $\eta_p^2 = .241$, a significant main effect of recall requirement, $F(3, 69) = 23.11$, $p < .001$, $\eta_p^2 = .501$, and a significant interaction effect between list length and recall requirement, $F(6, 138) = 7.33$, $p < .001$, $\eta_p^2 = .242$. Simple main effects revealed that for list length 4, the "recall all" condition was significantly different from the "recall 1" and "recall 3" conditions (ps at least $< .05$). For list lengths 5 and 6, the "recall 3" condition was significantly different from all other recall conditions (all ps at least $< .05$). Simple main effects also revealed that for the "recall 3" condition, list length 4 was significantly different from list lengths 5 and 6 (ps at least $< .05$). Figure 6D shows that the PFR was smaller for list length 4 than for the other two list lengths. Finally, for the "recall all" condition, list lengths 4

and 6 were significantly different from each other ($p < .01$). Figure 6D indicates that the PFR was greater for list length 6 than for list length 4.

Discussion

As with Experiments 1 and 2, it would appear from these findings that the recall demands of the task have had a large effect on the position in the list from which participants choose to initiate their recall. This effect is consistent across the three list length conditions of Experiment 3.

In line with previous research (Ward et al., 2010), participants consistently choose to recall the first item first when instructed to recall all the items, and this tendency is reduced as list length increases. Critically, when asked for recall of a single item, there is a tendency to choose to begin recall with one of the last few items; when asked to recall two items, participants are likely to begin with the penultimate item, and there is even a tendency for participants to begin with the antepenultimate item when asked to recall three items. In addition, for all three list lengths, the tendency to begin recall with the first item decreases steadily from the "recall all" to the "recall 1" conditions; however, the tendency to recall the last item first in the "recall 1" condition remains relatively constant across list lengths.

General Discussion

The findings from all three experiments are clear and consistent. Replicating recent work on the recall of short lists (e.g., Grenfell-Essam & Ward, 2012; Ward et al., 2010), we have shown that participants are most likely to initiate recall of short lists with the first list item when they are instructed to recall all the items in the list. However, a significant new finding is that this tendency appears to be driven by the requirement to recall as many list items as possible. Critically, when the number of words to be recalled is limited to one or two

items, participants are more likely to select a different recall strategy: they begin recall with one of the last few list items. We note that participants who were required to recall only one or two items often initiated their recall with the last or penultimate items, respectively, despite the fact that the list length was to some extent uncertain; this suggests that participants can to some extent choose the initial word to recall at retrieval.

We have previously argued that participants in immediate recall tasks can exert some control over which words, if currently accessible, they are to recall first (see, e.g., Bhatarah et al., 2008; Tan & Ward, 2007). Although increased rehearsal and increased attention can heighten the accessibility of certain list items at test (e.g., Bhatarah et al., 2009; Grenfell-Essam & Ward, 2015; Tan & Ward, 2000, 2008), under more normal circumstances, participants appear to have privileged access to the first list item and the recency items; the tendency to output the former decreases as its accessibility decreases with increasing list length (Ward et al., 2010).

This privileged access to the first list item may reflect the use of a start-of-list context cue (e.g., Davelaar, Goshen-Gottstein, Ashkenazi, Haarman, & Usher, 2005; Metcalfe & Murdock, 1981) or a "Get Ready" warning signal (e.g., Laming, 1999, 2010), or may reflect the heightened accessibility to the first list item within a group and/or list (e.g., Farrell, 2012). The privileged access to the recency items may reflect the output of a short-term store in dual store accounts of IFR (Anderson, et al., 1998; Davelaar et al., 2005; Raaijmakers & Shiffrin, 1981), or be the result of greater temporal distinctiveness (Brown et al., 2007) or a greater match with the end-of-list context (Howard & Kahana, 2002; Polyn, et al., 2009; Sederberg, et al., 2008; Tan & Ward, 2000) in unitary accounts of IFR.

Our interpretation of our findings is that even on relatively short lists of 4 to 6 words, participants have greater accessibility to the end-of-list items than they have to the first list item (despite its privileged status), but they nevertheless strive to recall the first item first in

immediate memory tasks when (1) this is a task requirement, such as in ISR, and (2) when they are to try to recall as many list items as possible, such as in the three immediate tasks reported here: ISR, IFR and ISR-free.

The advantage of retrieving the first list item first is that a forward-ordered recall strategy can facilitate the retrieval of multiple responses. Retrieval is often assumed to be both self-propagating and self-limiting (e.g., Roediger, 1973, 1974), such that the recall of one list item can facilitate the recall of the next (e.g., Howard & Kahana, 1999; Kahana, 1996; Nairne, Ceo & Reysen, 2007; see also Lohnas & Kahana, 2014), but can also cause output interference (e.g., Beaman, 2002; Bunting, Cowan & Saults, 2006, Cowan, Saults, Elliott & Moreno, 2002; Nairne et al., 2007; Oberauer, 2003; Tan & Ward, 2007). It may be that the "ISR-like" recall of short lists is an effective strategy to recall many of the words from short lists, but that when only one or two responses are required, participants favor the greater certainty of accessing only the most recent items.

Our interpretation of our data may go some way towards overcoming the difficulties that unitary accounts of IFR (e.g., Brown et al., 2007; Howard & Kahana, 2002; Polyn, et al., 2009; Sederberg, et al., 2008; Tan & Ward, 2000) have in explaining the IFR of short lists. Unitary accounts of IFR tend to predict extended recency effects, and so to date they have been found wanting in explaining participants' tendencies to initiate the recall of short lists with the first list item and to continue to recall in an "ISR-like" manner. However, our data suggest that, even in the IFR of very short lists, participants prefer to initiate recall with one of the last list items if they only have to recall one or two list items. This finding is entirely consistent with recency-based accounts of episodic memory and indicates that the tendency to initiate recall with the first item in a short list, leading to "ballistic" forward-ordered recall, is not obligatory; it also suggests that primacy-based mechanisms need not entirely account for the immediate recall of short lists (see also Tan & Ward, 2007).

We note that our findings further suggest that theoretical integration of immediate memory tasks may be possible (e.g., Anderson, et al., 1998; Bhatarah, et al., 2006, 2008; Brown, et al., 2007, 2008; Farrell, 2012; Grenfell-Essam & Ward, 2012; Grossberg & Pearson, 2008; Kahana, 2012; Klein, et al., 2005; Ward, et al., 2010). Consistent with recent comparisons, we observe similar patterns of new and established findings in IFR, ISR and ISR-free tasks when the methodologies, list lengths and scoring systems used in these immediate memory tasks are equated.

Finally, we note that the observed changes in recall strategy with different numbers of words to be recalled may potentially help reconcile the bowed serial position curves obtained in many immediate recall tasks (IFR, ISR, ISR-free, etc.) - when participants are requested to recall all the list items - with the more recency-based serial position curves obtained in single-probed or yes-no recognition tasks, when participants need only make a single response (e.g., Avons et al., 1994; Duncan & Murdock, 2000; Penney, 1982). As Duncan and Murdock (2000) suggest, there appears to be a rather different encoding strategy that occurs when participants anticipate trying to recall all or many of the list items compared to when they make only a single response. This is compatible with the finding by Oberauer (2003), who showed similar bowed serial position curves in serial recall and serial recognition tasks when both required the same number of multiple responses.

In summary, our three experiments have shown that the tendency to initiate recall of short lists with the first list item occurs only when participants are required to recall as many items from the list as they can. When participants are asked to recall only one or two items, they tend to initiate recall with end-of-list items. These findings show for the first time a manipulation that eliminates the initial tendency to recall in forward order, provide some support for recency-based accounts of IFR and help explain differences between single-response and multiple-response immediate memory tasks.

Acknowledgements

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.

Avons, S.E., Wright, K.L., & Pammer, K. (1994). The word-length effect in probed and serial recall. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 47,* 207-231.

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

Beaman, C. P. (2002). Inverting the modality effect in serial recall. *Quarterly Journal of Experimental Psychology, 55A*, 371-389.

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

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.

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.

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, 36*, 689-713.

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.

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

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

Bunting, M.F., Cowan, N., & Saults, J.S. (2006). How does running memory span work? *Quarterly Journal of Experimental Psychology, 59(10),* 1691-1700.

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

Corballis, M.C. (1967). Serial order in recognition and recall. *Journal of Experimental Psychology, 74(1)*, 99-105.

Cortis, C., Dent, K., Kennett, S., & Ward, G. (2015). First things first: Similar list length and output order effects for verbal and nonverbal stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1179-1214.

Cowan, N., Saults, J.S., Elliott, E.M., & Moreno, M. (2002). Deconfounding serial recall. *Journal of Memory and Language, 46*, 153-177.

Crowder, R.G. (1969). Behavioral strategies in immediate memory. *Journal of Verbal Learning and Verbal Behavior, 8*, 524-528.

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(1)*, 3-42.

Deese, J., & Kaufman, R.A. (1957). Serial effects in recall of unorganized and sequentially organized verbal material. *Journal of Experimental Psychology, 54,* 180-187.

Duncan, M., & Murdock, B. (2000). Recognition and Recall with Precuing and Postcuing. *Journal of Memory and Language 42,* 301–313.

Farrell, S. (2012). Temporal Clustering and Sequencing in Short-Term Memory and

    Episodic Memory. *Psychological Review, 119,* 223–271.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in

    serial recall. *Psychonomic Bulletin & Review, 9,* 59-79.

Glenberg, A.M., & Swanson, N.G. (1986). A temporal distinctiveness theory of recency and

    modality effects. *Journal of Experimental Psychology: Learning, Memory and*

    Cognition, 12, 3-15.

Golomb, J. D., Peelle, J. E., Addis, K. M., Kahana, M. J., and Wingfield, A. (2008). Effects

    of adult aging on utilization of temporal and semantic associations during free and

    serial recall. *Memory & Cognition*, *36(5)*, 947–956.

Grenfell-Essam, R., & Ward, G. (2012). Examining the relationship between free recall and

    immediate serial recall: The role of list length, strategy use, and test expectancy.

    Journal of Memory and Language, 67, 106-148.

Grenfell-Essam, R., & Ward, G. (2015). The effect of selective attention and a stimulus

    prefix on the output order of immediate free recall of short and long lists. Canadian

    Journal of Experimental Psychology, *69,* 1-16.

Grenfell-Essam, R., & Ward, G., & Tan, L. (2013). The role of rehearsal on the output order

    of immediate free recall of short and long lists. *Journal of Experimental Psychology:*

    Learning, Memory, and Cognition, 39, 317-347.

Grossberg, S. and Pearson, L. (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

Howard, M. W. and Kahana, M. J. (1999). Contextual variability and serial position effects

    in free recall. *Journal of Experimental Psychology: Learning, Memory, and*

    Cognition, *25*, 923–941.

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

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

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

Kahana (2012). Foundations of Human Memory. Oxford University Press.

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

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

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

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

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

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

Lohnas, L. J. and Kahana, M. J. (2014a). Compound cuing in free recall. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 40, 12-24.

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

Nairne, J. S., Ceo, D. A., & Reysen, M. B. (2007). The mnemonic effects of recall on immediate retention. *Memory & Cognition,* 35, 191-199.

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

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

Oberauer, K. (2003). Understanding serial position curves in short-term recognition and recall. *Journal of Memory and Language, 49,* 469–483.

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

Penney, C.G. (1982). Suffix effects and probe modality in probed recall: Implications for readout from sensory memory. *The Quarterly Journal of Experimental Psychology, 34A,* 245-257.

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

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

Roediger, H. L. (1973). Inhibition in recall from cueing with recall targets. *Journal of Verbal Learning and Verbal Behavior, 12*, 261-269.

Roediger, H. L. (1974). Inhibiting effects of recall. *Memory & Cognition, 2*, 261-269.

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

Spurgeon, J., Ward, G., & Matthews, W.J. (2014a). Why do participants initiate free recall of short lists with the first list item? Towa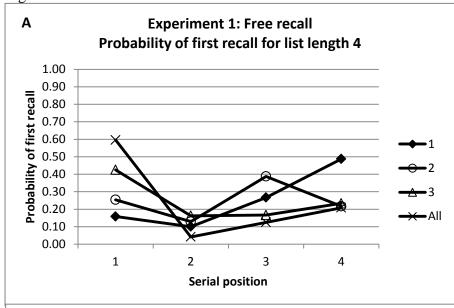rds a general episodic memory explanation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1551-1567.
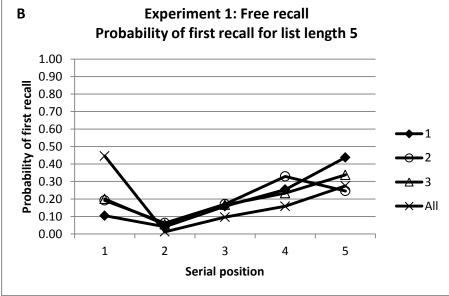
Spurgeon, J., Ward, G., & Matthews, W.J. (2014b). Examining the relationship between immediate serial recall and immediate free recall: Common effects of phonological loop variables, but only limited evidence for the phonological loop. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 1110-1141.

Spurgeon, J., Ward, G., Matthews, W.J., & Farrell, S. (2015). Can the effects of temporal grouping explain the similarities and differences between free recall and serial recall? *Memory & Cognition, 43*(3), 469-488.

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.

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

Tan, L. & Ward, G. (2008). Rehearsal in immediate serial recall. *Psychonomic Bulletin & Review, 15*, 535-542.

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

Waugh, N.C. (1960). Serial position and the memory-span. *American Journal of Psychology, 73*, 68-79.

<u>Figure Captions</u>

<u>Figure 1.</u> Data from Experiment 1 showing the probability of first recall as a function of

serial position for list lengths 4 (Figure 1A), 5 (Figure 1B) and 6 (Figure 1C).

<u>Figure 2.</u> Data from Experiment 1 showing the probability of first recall as a function of list

length for the first (Figure 2A), final (Figure 2B), penultimate (Figure 2C) and

antepenultimate (Figure 2D) serial positions.

<u>Figure 3.</u> Data from Experiment 2 showing the probability of first recall as a function of

serial position for list lengths 4 (Figure 3A), 5 (Figure 3B) and 6 (Figure 3C).

<u>Figure 4.</u> Data from Experiment 2 showing the probability of first recall as a function of list

length for the first (Figure 4A), final (Figure 4B), penultimate (Figure 4C) and

antepenultimate (Figure 4D) serial positions.

<u>Figure 5.</u> Data from Experiment 3 showing the probability of first recall as a function of

serial position for list lengths 4 (Figure 5A), 5 (Figure 5B) and 6 (Figure 5C).

<u>Figure 6.</u> Data from Experiment 3 showing the probability of first recall as a function of list

length for the first (Figure 6A), final (Figure 6B), penultimate (Figure 6C) and

antepenultimate (Figure 6D) serial positions.

Figure 1

A **Experiment 1: Free recall**
Probability of first recall for list length 4

B **Experiment 1: Free recall**
Probability of first recall for list length 5

C **Experiment 1: Free recall**
Probability of first recall for list length 6

Figure 2

A Experiment 1: Free recall
Probability of first recall (PFR) for the first serial position

B Experiment 1: Free recall
Probability of first recall (PFR) for the final serial position

C Experiment 1: Free recall
Probability of first recall (PFR) for the penultimate serial position

D Experiment 1: Free recall
Probability of first recall (PFR) for the antepenultimate serial position

Figure 3

A

Experiment 2: ISR
Probability of first recall for list length 4

B

Experiment 2: ISR
Probability of first recall for list length 5

C

Experiment 2: ISR
Probability of first recall for list length 6

Figure 4

A
Experiment 2: ISR
Probability of first recall (PFR) for the first serial position

B
Experiment 2: ISR
Probability of first recall (PFR) for the final serial position

C
Experiment 2: ISR
Probability of first recall (PFR) for the penultimate serial position

D
Experiment 2: ISR
Probability of first recall (PFR) for the antepenultimate serial position

Figure 5

A

Experiment 3: ISR-free
Probability of first recall for list length 4

Probability of first recall

Serial position

1
2
3
All

B

Experiment 3: ISR-free
Probability of first recall for list length 5

Probability of first recall

Serial position

1
2
3
All

C

Experiment 3: ISR-free
Probability of first recall for list length 6

Probability of first recall

Serial position

1
2
3
All

Figure 6

A   Experiment 3: ISR-free
Probability of first recall (PFR) for the first serial position

B   Experiment 3: ISR-free
Probability of first recall (PFR) for the final serial position

C   Experiment 3: ISR-free
Probability of first recall (PFR) for the penultimate serial position

D   Experiment 3: ISR-free
Probability of first recall (PFR) for the antepenultimate serial position