Memory, 2013
Vol. 21, No. 3, 347–365, http://dx.doi.org/10.1080/09658211.2012.726629

Routledge
Taylor & Francis Group

Serial-position effects for items and relations in short-term memory

Tim Jones[1] and Klaus Oberauer[1,2]

[1]Department of Psychology, University of Bristol, Bristol, UK
[2]Department of Psychology - Cognitive Psychology Unit, University of Zurich, Switzerland

Two experiments used immediate probed recall of words to investigate serial-position effects. Item memory was tested through probing with a semantic category. Relation memory was tested through probing with the word's spatial location of presentation. Input order and output order were deconfounded by presenting and probing items in different orders. Primacy and recency effects over input position were found for both item memory and relation memory. Both item and relation memory declined over output position. The finding of a U-shaped input position function for item memory rules out an explanation purely in terms of positional confusions (e.g., edge effects). Either these serial-position effects arise from variations in the intrinsic memory strength of the items, or they arise from variations in the strength of item–position bindings, together with retrieval by scanning.

Keywords: Serial-position curves; Short-term memory; Primacy effects; Recency effects; Computational modelling.

For over a hundred years, researchers have investigated people's performance on simple memory tasks to better understand the cognitive processes involved in remembering and forgetting (e.g., Ebbinghaus, 1885/1964; Nipher, 1876, 1878). One such task is to present people with a list of items and then ask them to recall the items immediately after presentation. The results tend to show a U-shaped *serial-position curve*, with better memory for items presented early in the list (primacy effect) and late in the list (recency effect) than for those in the middle.

Although the bow-shaped serial-position curve is one of the oldest and best established facts about short-term memory, no comprehensive theory has yet emerged explaining it. There are, however, a number of proposed mechanisms that could cause primacy and recency effects. The goal of our present work is to contribute further evidence for narrowing down the set of candidate mechanisms. In particular, we aim to distinguish serial-position effects on two different tests of memory: item memory and relation memory. Item memory refers to remembering which items have been presented on a list, irrespective of their positions. Relation memory additionally requires memory for relations between items (e.g., their order in the list), or memory for relations of items to some aspect of their context (e.g., which item has been presented in which location). The distinction between item and relation memory is theoretically interesting because it is possible to maintain items purely through the temporarily sustained activation of their representations,

Address correspondence to: Tim Jones, Department of Psychology, University of Bristol, 12a Priory Road, Bristol, BS8 1TU, UK. E-mail: Timothy.Jones@bristol.ac.uk or k.oberauer@psychologie.uzh.ch

This research was supported by a grant from the Economic and Social Research Council (ESRC), Grant RES-062-23-1199. Klaus Oberauer is now at the Department of Psychology–Cognitive Psychology Unit, University of Zurich, Switzerland.

whilst relation memory arguably requires the binding of each item to a context or position marker (e.g., Brown, Neath, & Chater, 2007; Brown, Preece, & Hulme, 2000; Burgess & Hitch, 1999, 2006; Farrell, 2006, 2012; Henson, 1998). Therefore, disentangling effects of serial position on item memory and on relation memory provides information on whether serial-position affects the memory strength of items (e.g., through sustained activation) or the strength of item–context bindings, or both.

Input order and output order

Many common methodologies for investigating short-term memory confound the order in which items are encoded (input order) with the order in which they are retrieved (output order). For example, in forward recall, output order is always the same as input order, and the resulting curve tends to show large primacy and only a little recency, whereas in backward recall, output order is always the reverse of input order, and the resulting curve typically shows large recency and only a little primacy (e.g., Li & Lewandowsky, 1993, 1995; Madigan, 1971). This demonstrates the importance of output order in shaping serial-position curves. Effects of input order and output order need to be separated because they most likely reflect different mechanisms. Effects of input order are likely to arise from processes during encoding, whereas effects of output order must arise from processes during retrieval. Experiments investigating the effects of input order and output order separately (Cowan, Saults, Elliott, & Moreno, 2002; Oberauer, 2003) reveal that they have different effects on memory performance. In this article, we focus on the effects of input serial position for two reasons: First, there is a richer set of theoretical ideas about the sources of input-position effects (see Table 1) than for those of output-position effects (for a review of hypothetical sources of both kinds of effects, see Oberauer, 2003). Second, input-position effects are more general because memory for lists always involves input order, but output order comes into play only when more than one item is tested.

Several methods can be used to control for effects of output order. One is to test for only one item per trial (e.g., Davelaar, Goshen-Gottstein, Ashkenazi, Haarman, & Usher, 2005; Hay, Smyth, Hitch, & Horton, 2007; McElree & Dosher, 1989; Monsell, 1978). Another method is to test all of the items from a trial, one at a time,

TABLE 1
Potential loci of primacy and recency, and the predictions of the generic model for item and relation memory

Type of memory	*Locus of serial-position effects*	*Retrieval mechanism*	*Serial-position curve*
Item memory	**Item activation**	Direct cueing with category	U
	Item–position bindings		Flat
	Positional distinctiveness		Flat
	Item activation	Scanning input positions	U
	Item–position bindings		U
	Positional distinctiveness		Inverted-U
Relation memory	**Item activation**	Direct cueing with spatial location	U
	Item–position bindings		Flat
	Positional distinctiveness		Flat
	Item–location bindings		U[a]
	Position–location bindings		Flat
	Spatial distinctiveness		Flat
	Item activation	Scanning input positions[b]	U
	Item–position bindings		U
	Positional distinctiveness		U
	Item–location bindings		Flat
	Position–location bindings		U
	Spatial distinctiveness		U

Note. Bold font indicates viable combinations of mechanisms according to our data. Position refers to serial input position. Location refers to spatial location. U = U-shaped serial-position curve with primacy and recency. Flat = flat serial-position curve with no primacy and no recency. Inverted-U = upside down serial-position curve, with better memory for middle positions.
[a]Only if bindings are stronger at extreme input positions, not extreme spatial locations. [b]In all cases the scanning retrieval mechanism shifts the serial-position curves a little towards primacy.

in a different order from their order of presentation, such that input order and output order are uncorrelated (Cowan et al., 2002; Oberauer, 2003). In the present experiments we used the second method, for reasons that are explained below.

Serial-position effects in tests of item memory and relation memory

Previous experiments testing input serial-position effects suggest that item memory is characterised by a pronounced recency advantage, with little or no primacy effect, whereas relation memory demonstrates a more bow-shaped curve. A measure of item memory can be obtained by tests of item recognition that ask whether a probe stimulus has been presented in the memory list, regardless of its relation to other list items or to a particular position in the list. When a single item is probed shortly after encoding a memory list, serial-position curves typically show extended recency and only a small, one-item primacy effect (e.g., Hay et al., 2007; McElree & Dosher, 1989; Monsell, 1978). When item recognition is tested by a series of probes rather than a single probe, such that test order is unconfounded with presentation order, the recency effect is still larger than the primacy effect (Oberauer, 2003).

Item memory can also be tested by probed recall using cues that are already related to the items through a permanent association in long-term memory, rather than through a temporary association formed on the current trial. For instance, Davelaar et al. (2005) presented a list of words that all belonged to different categories and then probed for a particular item by presenting its category as a cue. This is the method used to test item memory in the current study. The semantic links between items and categories are established in long-term memory, so that the correct response can be determined from the long-term semantic association between category and item, together with memory for which items have been presented on the current list. Memory for the relations of the current list items to each other, or to their list positions, are irrelevant for performance. Davelaar et al. (2005) found clear recency and no primacy when probing for single items from a list by category name at typical presentation rates (800 ms per item); primacy emerged only at very rapid rates (300 ms per item or faster). These findings suggest that, at least with conventional presentation rates, tests of item memory result in a strong recency effect, with little or no primacy effect.

Relation memory can be tested with a single probe indicating the requested item's serial position (e.g., Farrell & Lelièvre, 2009; Nairne, Ceo, & Reysen, 2007), or the spatial location associated with the requested item (e.g., Avons, Wright, & Pammer, 1994). In either case, the result tends to be a bow-shaped serial-position curve, with more pronounced primacy than for item memory.

This comparison of findings in the literature suggests that the recency effect reflects primarily an advantage in item memory (activation strength) for the more recently presented items, whereas the primacy effect reflects an advantage in relation memory (item–context bindings) for the initially presented list items. However, this comparison is compromised by two challenges, which we address next.

How to separate and compare item memory and relation memory

An effort to separately assess serial-position effects on item memory and relation memory faces two challenges. The first challenge is that tests of item memory and tests of relation memory often differ not only in whether they require item memory or relation memory, but also in other regards, which confound their comparison. Such potential confounding variables include the method of testing (recall vs. recognition), the number of items tested (a single item or the entire list), and people's expectancy (i.e., knowing in advance whether item memory or relation memory will be tested). These variables are likely to affect the shape of the serial-position curve (e.g., Bhatarah, Ward, & Tan, 2008; Duncan & Murdock, 2000; Oberauer, 2003).

The second challenge arises from the task-purity problem: We cannot be certain that tests of relation memory actually draw on relational representations (i.e., bindings), and we cannot be certain that tests of item memory do not rely on relational representations. For instance, a test of relation memory that has been much studied in the short-term memory literature is serial recall, requiring memory for items in their order of presentation. Order memory is explained in many models of serial recall by bindings between each item and a position marker (e.g., Brown et al., 2000; Burgess & Hitch, 1999, 2006; Farrell, 2006, 2012; Henson, 1998). However, forward serial recall can also be accomplished without such bindings, by establishing an activation gradient across the item representations, such that successive list items are activated to a decreasing level (Grossberg & Stone, 1986; Page & Norris, 1998). Conversely, tests of item memory might not rely exclusively on the memory strength of items. For example, one way to measure item memory is to request that participants recall items in their appropriate positions, but then score an item as correct even if it is produced in the wrong position, as long as it was on the list (e.g., Bjork & Healy, 1974; Fuchs, 1969). With this method, recall is almost certainly dependent on memory for item–position bindings.

In the present study we address these challenges as follows: We tested both item memory and relation memory by probed recall. Participants encoded a list of words presented in random order across a vertical array of spatial locations. Each word belonged to a different semantic category. Relation memory was tested by probing every word from the list with its spatial location in a new random order (Oberauer, 2003). Recalling an item by its spatial location requires memory for the relation between each word and its spatial location. Item memory was tested by probing every word from the list with its semantic category in a new random order (Davelaar et al., 2005). This test requires memory for which words were presented in the current list, because over the entire experiment several words belonging to each category are presented, but it requires no memory for any relations of the current list words to other words or to their current context. The relation between a category probe and its associated list word is part of semantic long-term memory and therefore does not have to be held in short-term or working memory.

This experimental paradigm equates tests of item and of relation memory for testing method (i.e., probed recall) and number of items tested (i.e., the entire list). To control test expectancy, in Experiment 1 we included a postcued condition in which the kind of probe (spatial probes or category probes) was revealed only after encoding (Duncan & Murdock, 2000). By randomising input order and output order separately on each trial, we deconfound the effects of input and of output order (Oberauer, 2003); in this article we concentrate on the effect of input serial position.

Presentation of items in random order across the spatial locations also ensures that input order is uncorrelated with spatial location. For this reason people cannot easily rely on a primacy gradient over input order to recover which item was in which spatial location. Therefore, our test of relation memory arguably tests temporary bindings between words and their spatial locations. Conversely, our test of item memory is arguably a fairly pure test of item memory, because people's score depends only on whether or not they remember which member of the probe category was presented on the current list. It is not necessary to recover the bindings between a word and its input position or spatial location to achieve item recall.

These efforts to address the task-purity problem notwithstanding, we cannot be certain that recall of the item in a given spatial location actually relies on item–location bindings, or that recall of the item belonging to a given category does not rely on any item–context bindings. Therefore, we do not rely on our two kinds of test as pure tests of item memory and of relation memory, respectively. Rather, we consider a range of possible retrieval processes through which people could perform these two tests and implement these processes in computational models. The models combine these alternative retrieval processes with various mechanisms that have been proposed as explanations for serial-position effects at encoding. We turn to these proposed mechanisms next.

Proposed causes of serial-position effects at encoding

There are several proposed causes of primacy and recency effects over input position in immediate memory. The current study aims to constrain the broad locus where serial-position effects occur, rather than specify their exact mechanistic causes. For example, serial-position effects might be established over the *strength of item representations*; over the *strength of bindings* between item and context representations (i.e., bindings between items and input-position markers, or between items and spatial locations); or through the *distinctiveness* of context representations. A factorial combination of these possibilities with various retrieval mechanisms leads to a large number of potential sources of serial-position effects (see Table 1). We believe that rejecting any implausible combinations is a valuable step in narrowing down the potential mechanisms involved. Here, the potential causes and retrieval processes considered are outlined. Following presentation of the empirical results, we report the outcome of modelling all possible combinations of loci of serial-position effects with different retrieval mechanisms and rule out certain combinations of mechanisms as implausible.

Strength of item representations. One potential cause of serial-position effects is the strength (e.g., activation) of item representations. Item strength could vary through the amount of attention paid to each item (e.g., Farrell & Lewandowsky, 2002; Page & Norris, 1998). It has often been assumed that more attention is paid to early items in a list (e.g., Brown et al., 2000), possibly due to their novelty (e.g., Farrell & Lewandowsky, 2002), creating a primacy gradient in item strength, which should translate into a primacy effect in both tests of item memory and tests of relation memory. A recency gradient in item strength could arise from retroactive interference during encoding. This occurs when the encoding of later items interferes with the maintenance of earlier items, perhaps because later items steal features from earlier items (e.g., Nairne, 1990; Oberauer & Kliegl, 2006). A recency gradient in item strength should generate a recency effect in tests of item and of relation memory.

Strength of item–context bindings. The strength of item–context bindings may also be affected by attention during encoding. The oscillator-based associative recall model (OSCAR; Brown et al., 2000) assumes a gradual decline over input position in the learning rate for associations between items and position markers. This would lead to a primacy gradient if and only if items are retrieved through associations with their input-position markers. The strength of item–context bindings could also form a recency gradient (i.e., stronger bindings towards the end of the list), if retroactive interference impairs item–context bindings (Oberauer & Kliegl, 2006). In the present paradigm, a gradient of binding strength can apply either to bindings between items and markers of their input position, or to bindings between items and their spatial locations.

Contextual distinctiveness. In the literature on immediate serial recall, primacy and recency effects are often explained through the different *distinctiveness* of their contexts. These distinctiveness mechanisms assume that items are retrieved by cueing with their associated context markers. For instance, if items are cued by position markers, the *positional distinctiveness* of these markers will affect the likelihood of order errors at each input position. In its simplest form, this kind of explanation appeals to *edge effects*: The first and last items in a list have only one neighbouring item along the input position dimension, whereas items in the middle of the list have two neighbours. This means that the first and last items have less chance of having their position confused with neighbouring items (e.g., Brown et al., 2007; Brown et al., 2000; Henson, Norris, Page, & Baddeley, 1996; Lee & Estes, 1977).

A more sophisticated version of positional distinctiveness is implemented in Henson's (1998) start–end model. The positional context is formed by a combination of two activation gradients: a "start" gradient, with activation decreasing exponentially from the start of the list, and an "end" gradient, with activation decreasing exponentially from the end of the list. Recall is initiated by reinstating the context for the item at a certain position. Due to the exponential shape of the gradients, the context is most distinctive from other contexts at the start and the end of the list. Therefore, there is less chance of confusion between these items and others with neighbouring contexts, leading to better recall for start and end items (primacy and recency effects). Henson (1998) suggested that start–end context markers do not have to denote input position, but could also demarcate other relevant contexts (e.g., spatial locations demonstrating *spatial distinctiveness*), if these are more suitable for the task. As our test of relation memory involves probing with spatial location cues, we consider the effects of distinctiveness across both the input position and the spatial location in our computational models.

A further variant of positional distinctiveness has been proposed by Botvinick and Watanabe (2007). In their model of serial recall, representations of numerical rank serve as position markers. Based on single-cell recordings of rank-sensitive neurons in the parietal cortex, they assumed that the distinctiveness of rank representations decreased with increasing numerical rank. This translates into a primacy gradient on distinctiveness of position markers for input position.

Retrieval mechanisms. We consider two mechanisms by which items are retrieved in tests of item memory. The first is when the temporary activation of items above a threshold level discriminates them from less activated representations of other items (e.g., Farrell & Lewandowsky, 2002; Page & Norris, 1998). In this case, only variability in item activation levels could cause serial-position effects. Alternatively, item memory can be established by binding all items of the current trial to a context representation that distinguishes the current trial from preceding trials (e.g., Brown et al., 2007; Burgess & Hitch, 2006; Henson, 1998; Lewandowsky & Farrell, 2008). The context representation is then used to cue the associated items, and the strength of binding to items can also contribute to serial-position effects.

The most obvious mechanism for remembering the relation between an item and its spatial location over the short term is by maintaining direct bindings between item representations and representations of corresponding locations. In this case, the strength of item–location bindings, and the spatial distinctiveness of location representations, would determine recall. However, it is possible that both items and locations are bound to position markers coding their input order, and people scan through the input-position markers to probe for both items and their locations (e.g., Cowan, Saults, & Morey, 2006). In this case, any direct binding between spatial locations and item representations would be irrelevant for retrieval. Serial-position curves would be determined instead by the strength of bindings between input-position markers and items, and between input-position markers and spatial locations, as well as the distinctiveness of the input-position markers.

The present experiments

Our study aims to constrain the possible mechanisms that contribute to serial-position effects when encoding a list of items into short-term memory. To this end, we provide a controlled empirical separation of tests of item and relation memory and compare the results to the predictions of computational models of various potential memory mechanisms. We used Oberauer's (2003) random probed recall method for deconfounding input and output order. The method used for tests of item and relation memory were identical until retrieval. Category cues provide a test of item memory, whereas spatial location cues test relation memory.

We report two experiments: Experiment 1 aimed to investigate the effects of probe type (i.e., probing item memory with category cues, and probing relation memory with location cues) on primacy and recency effects over input position. In addition, the impact of strategic encoding was investigated by including a manipulation of pre- versus postcueing. To foreshadow, we obtained U-shaped serial-position curves with both primacy and recency for tests of item memory as well as tests of relation memory, different from Davelaar et al. (2005). Experiment 2 examined the impact of proactive interference on primacy and recency over input position to test one possible cause for the difference between our serial-position curves for item memory and those of Davelaar et al.

EXPERIMENT 1: PROBE TYPE AND STRATEGIC ENCODING

The goal of Experiment 1 was to separate the contributions of item and relation memory to primacy and recency effects over input position on memory for lists. We used category probes to test item memory and spatial-location probes to test relation memory. To isolate serial-position effects of input position from the confounding effects of output order, we tested items in a new, random order that was uncorrelated with input order.

To explore possible effects of strategic encoding based on task demands (e.g., Duncan & Murdock, 2000), participants learnt whether they would be tested by a category probe or a spatial probe either before list presentation (precueing) or after list presentation (postcueing).

We expected that item memory would show mostly recency, because this is what previous investigations have found using category probes (Davelaar et al., 2005) or item recognition (Monsell, 1978; Oberauer, 2003). For relation memory (positional probes) we expected primacy and recency effects over input position of about equal size, or even more primacy than recency, based on previous findings from serial recall, probed recall, and relational recognition (e.g., Cowan et al., 2002; Farrell & Lelièvre, 2009; Healy, 1974; Oberauer, 2003). As noted in the introduction, the comparison of item memory and relation memory across these experiments is compromised by various confounds. The present study provides a direct, within-experiment comparison of a test of item memory and a test of relation memory avoiding these confounds.

Method

Participants. Sixteen participants, 15 women and 1 man between the ages of 17 and 35 years, took part in Experiment 1. They were all psychology undergraduates at the University of Bristol who were native speakers of English and participated to receive course credit.

Design. Each trial included seven English words from seven different sets (described in the *Materials* section below), presented visually in seven different boxes arranged vertically on a computer screen. All seven words were then probed for recall using either item probes or spatial position probes. The probe type was indicated either before the trial began (precueing) or after the words had been presented (postcueing). The combination of pre/post cueing and probe type resulted in four within-subjects trial types. The spatial position of each word was determined at random on each trial. The order in which words were probed for recall (i.e., their output order) was also varied randomly on each trial, with the constraint that every combination of input position and output position (e.g., word at input position 3 probed for recall at output position 2) occurred four times for each trial type. For each trial type this required 196 (= $7 \times 7 \times 4$) words to be recalled, using 28 trials, resulting in a total of 112 trials. The type of cueing was manipulated between sessions. Participants took part in two sessions with 56 trials in each. Half of the participants took part in the precued trials in their first session, and the other half carried out the postcued trials in their first session. Within the precued trials for each session, half of the participants did the 28 category-probed trials first, whilst the other half did the 28 spatial-probed trials first. In the postcued session, probe types occurred in random order. Categories were chosen randomly for each trial, with no category repeated on consecutive trials. Words were chosen at random from 10 possibilities for each category. The first session was preceded by eight practice trials, and the second session was preceded by four practice trials, all of which were excluded from the data analysis.

Materials. Words of one or two syllables were selected from the category norms of Van Overshelde, Rawson, & Dunlosky (2004). We created 17 sets of 10 words belonging to 17 categories; an attempt was made to minimise the variation of word frequency within each category using frequencies from the CELEX database (Baayen, Piepenbrock, & van Rijn, 1995). Each trial was constructed by selecting seven sets at random, avoiding repetitions of sets across consecutive trials, and then selecting from each set one of the 10 words at random. Before the experiment began, the participant was presented with a list of all 170 words to be used in the experiment, grouped into the 17 categories, to clarify the category membership of the words.

Procedure. Each trial began with a cue to the trial type (‘‘CATEGORIES’’, ‘‘QUESTION MARKS’’, or ‘‘START’’) displayed in the middle of a computer screen using black text on a white background. Participants were cued with ‘‘categories’’ for the category-probed trials, ‘‘question marks’’ for the spatial-probed trials, and ‘‘start’’ for the postcued trials. After 2 s, this cue disappeared, and 500 ms later seven black rectangles appeared, one above the other, left of the centre of the screen. Following a 500-ms pause, the seven words were presented, one at a time, in different boxes in a random order. Each word remained visible for 800 ms; the offset of one word coincided with the onset of the next word. After the last word disappeared, and a further 700 ms pause, the first probe appeared. For spatial position probes, this involved the presentation of a red question mark in one of the boxes, in which case the participant was expected to say out loud the word that appeared in that box for the experimenter to write down. For category probes, a category name would appear in red to the right of the boxes, and the participant was expected to say out loud the word in the memory list belonging to that category. Participants were instructed to say ‘‘don’t know’’ if they couldn’t remember a word. Once the word had been recalled, the participant pressed the space bar, and after 200 ms the next probe appeared, and so on until all seven words had been recalled. After recalling the final word of the trial and pressing the space bar, there was a 3.5 s pause with a blank white screen, and then the next trial began with the display of the trial type cue.

Before each block of precued trials, an on-screen message made clear how the next 28 trials would be probed: with ‘‘categories’’ or ‘‘question marks’’. Participants were tested for two 60-minute sessions consisting of four blocks each, and they were encouraged to have a break after each block.

Results

An analysis of variance (ANOVA) was carried out with four within-subjects factors: pre/post cueing, probe type, input position, and output position, and proportion of answers correct in the appropriate position as the dependent variable. The results up to quadratic contrasts are presented in Table 2.

Figure 1 shows that there were roughly symmetrical primacy and recency effects over input position. The lack of an interaction with probe type indicates that the shape of the input position curve was similar for spatial probes and category probes. Serial position did interact with the type of cueing, however: Precueing resulted in an elevated primacy effect (see Figure 1). This observation is reflected in the significant main effect of cueing and the interaction of cueing with the linear contrast of input position. The lack of a three-way interaction between cueing, probe type, and input position suggests that the effect of precueing was the same for both probe types.

TABLE 2
Results of ANOVA for Experiment 1, testing effects of pre/postcueing, probe type, input position, and output position

Variables	*Contrast*	F	MSE	p	η_p^2
Pre/Post		5.48*	0.281	.033	.268
Probe type		142.52***	0.557	<.001	.905
Input	L	0.158	1.025	.697	.01
	Q	41.113***	0.19	<.001	.733
Output	L	147.889***	0.16	<.001	.908
	Q	72.573***	0.041	<.001	.829
Pre/Post × Input	L × L	6.834*	0.188	.02	.313
	L × Q	1.65	0.08	.218	.099
Probe Type × Input	L × L	3.415	0.072	.084	.185
	L × Q	2.306	0.134	.15	.133
Pre/Post × Output	L × L	0.038	0.102	.847	.003
	L × Q	0.202	0.073	.66	.013
Probe Type × Output	L × L	0.694	0.11	.418	.044
	L × Q	0.538	0.059	.474	.035
Output × Input	L × L	81.757***	0.025	<.001	.845
	L × Q	10.383**	0.08	.006	.409
	Q × L	17.914***	0.05	.001	.544
	Q × Q	5.516*	0.035	.033	.269
Pre/Post × Probe Type × Input	L × L × L	1.148	0.078	.301	.071
	L × L × Q	1.889	0.052	.189	.112
Pre/Post × Probe Type × Output	L × L × L	0.208	0.031	.665	.014
	L × L × Q	0.306	0.033	.588	.02
Pre/Post × Output × Input	L × L × L	0.114	0.03	.74	.008
	L × L × Q	1.242	0.041	.283	.076
	L × Q × L	3.861	0.035	.068	.205
	L × Q × Q	3.465	0.035	.082	.188
Probe Type × Output × Input	L × L × L	0.051	0.077	.825	.003
	L × L × Q	15.602***	0.045	.001	.51
	L × Q × L	0.002	0.057	.962	0
	L × Q × Q	4.55*	0.049	.05	.233
Pre/Post × Probe Type × Output × Input	L × L × L × L	1.037	0.068	.325	.065
	L × L × L × Q	0.173	0.06	.684	.011
	L × L × Q × L	4.627*	0.034	.048	.236
	L × L × Q × Q	0.913	0.05	.354	.057

Note. The degrees of freedom in all cases are (1, 15). ANOVA = analysis of variance. η_p^2 = partial eta squared; L = linear contrast; Q = quadratic contrast; L × Q = interaction of linear contrast and quadratic contrast; Pres. rate = presentation rate; Input = input position; Output = output position.
**p* <.05, ***p* <.01, ****p* <.001.

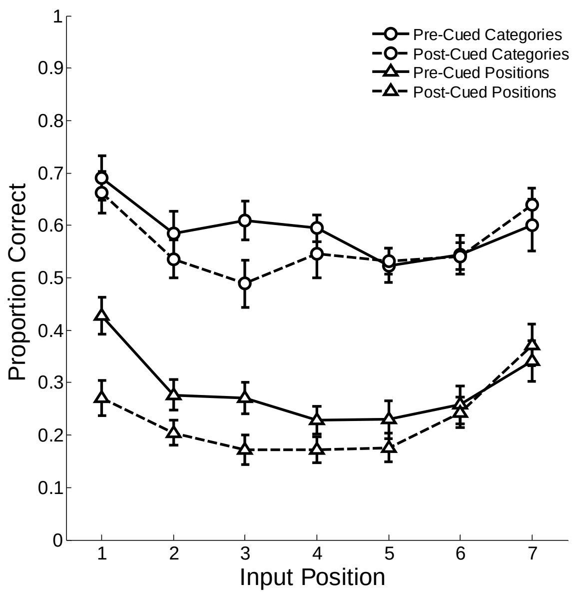

Figure 1. Accuracy as a function of probe type, pre/postcueing, and input position in Experiment 1 averaged over all output positions. Error bars represent within-subjects standard errors (e.g., Bakeman & McArthur, 1996).

There was a clear decline in accuracy over output position, and an interaction between input position and output position, with later input positions being more strongly affected by output interference.

The main effect of probe type reflected much better recall with category probes than positional probes. There were also some significant higher order interactions that are difficult to interpret, so are not discussed here.

For a more direct comparison with previous studies that probed only one item per trial (e.g., Davelaar et al., 2005), the effects of type of cueing, probe type, and input position were investigated for only the first output position. The significant contrasts from this ANOVA are reported in Table 3. The significant linear and quadratic contrasts over input position reflect an input position curve with both primacy and recency; the recency effect was more pronounced than the primacy effect (see Figure 2). The quadratic contrast of input position interacted with type of cueing and with probe type. These interactions reflect the fact that the middle items suffered most from postcueing (compared to precueing) and from using spatial location probes (compared to category probes).

Discussion

There were clear serial-position effects over input position for both item and relation memory, even though memory of input position is not directly required to complete either task. A comparison of serial-position curves over input position for item and relation memory showed no differences in primacy and recency when averaging over all output positions, but slightly greater primacy and recency for relation memory than item memory at the first output position. The first output position resulted in best recall, and item memory was generally superior to relation memory, so that the serial-position curve for item memory when only the first-probed item was considered was close to ceiling—this alone could have flattened this serial-position curve relative to that for relation

TABLE 3

Results of ANOVA for Experiment 1, testing effects of pre/postcueing, probe type, and input position at Output Position 1

Variables	*Contrast*	F	MSE	p	η_p^2
Pre/post		4.03	0.097	.063	.212
Probe type		137.627***	0.085	<.001	.902
Input pos.	L	10.257**	0.16	.006	.406
	Q	93.082***	0.04	<.001	.861
Pre/Post × Probe Type		0.352	0.067	.562	.023
Pre/Post × Input	L × L	2.119	0.061	.166	.124
	L × Q	12.954**	0.038	.003	.463
Probe Type × Input	L × L	3.064	0.042	.1	.17
	L × Q	14.887**	0.057	.002	.498
Pre/Post × Probe Type × Input	L × L × L	2.288	0.064	.151	.132
	L × L × Q	0.133	0.039	.72	.009

Note. The degrees of freedom in all cases are (1, 15). ANOVA = analysis of variance. η_p^2 = partial eta squared; L = linear contrast; Q = quadratic contrast; L × Q = interaction of linear contrast and quadratic contrast; Pre/post = pre/postcueing; Input = input position.
*$p < .05$, **$p < .01$, ***$p < .001$.

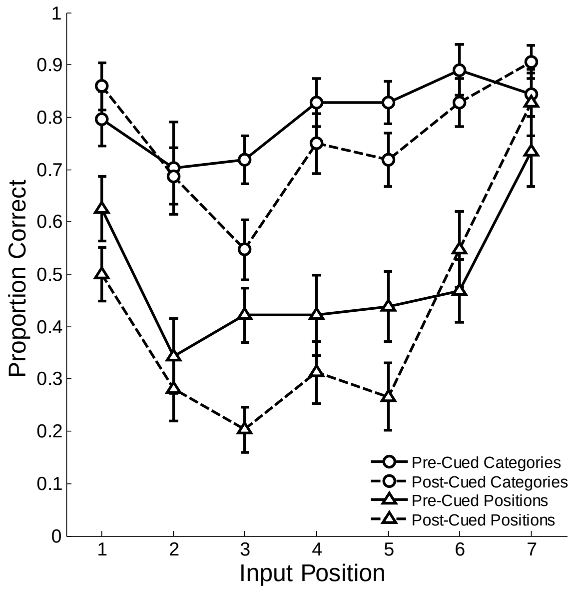

Figure 2. Accuracy as a function of probe type, pre/postcueing, and input position for Experiment 1 at only the first output position. Error bars represent within-subjects standard errors (e.g., Bakeman & McArthur, 1996).

memory. Therefore, the conservative conclusion at this point is that serial-position curves for item memory and for relation memory were roughly parallel.

The presence of a quadratic contrast over input position without a linear contrast indicates symmetrical primacy and recency for both item and relation memory. Previous experiments have found predominantly recency for tests of item memory (e.g., Hay et al., 2007; McElree & Dosher, 1989; Monsell, 1978; Oberauer, 2003), and Davelaar et al. (2005) found *only* recency with no primacy. Two methodological features might be held responsible for the difference between our results and those of previous studies. In previous studies participants had advance knowledge of the kind of probe, and tests of item memory often involved the output of a single item per trial (e.g., Davelaar et al., 2005; Hay et al., 2007; McElree & Dosher, 1989; Monsell, 1978).

The first explanation can be ruled out by our finding that the primacy effect was, if anything, even stronger in the precued condition than in the postcued condition. The second explanation appears more promising. When we limited the analysis to the first output position for a direct comparison to experiments testing only a single item, we obtained more pronounced recency than in the overall analysis. A small primacy effect was still found for both item-memory and relation-memory tests. This matches many previous tests of item memory (e.g., Hay et al., 2007; McElree & Dosher, 1989), although not Davelaar et al. (2005), whose methodology was most similar to ours and resulted in no measurable primacy effect. The difference between our results and those of Davelaar et al. is explored further in Experiment 2.

There was a general decline in accuracy with output position (Cowan et al., 2002; Oberauer, 2003), confirming a role of output interference in item and relation memory. The interaction between input position and output position suggests that output interference had different effects depending on the input position. Like Cowan et al. (2002), we found that output interference impacted on the recency portion of the serial-position curve more strongly than the primacy portion. This explains why the serial-position curve at the first output position, which is unaffected by output interference, shows less primacy and more recency than the serial-position curve averaged over all output positions.

To conclude, we obtained essentially parallel serial-position curves over input position for tests of item memory and of relation memory. This result goes against our expectation, gleaned from the literature, that item memory is characterised more by recency, whereas relation memory is characterised by serial-position curves that are symmetrically U-shaped, or even have stronger primacy than recency. Our results therefore do not support the speculation that the recency effect in tests requiring both item and relation memory is primarily driven by item memory, whereas the primacy effect is primarily driven by relation memory. We explore the theoretical implications of our findings more thoroughly through a systematic exploration of computational models, described after Experiment 2.

EXPERIMENT 2: PROACTIVE INTERFERENCE

Davelaar et al. (2005) suggested that long-term memory may contribute to the primacy effect in tests of short-term memory, perhaps because earlier items have a greater opportunity for rehearsal (e.g., Tan & Ward, 2008) and are therefore more likely to be stored into long-term memory (e.g., Rundus, 1971). In their Experiment 2, Davelaar et al. used a small pool of potential words for each category and repeated categories from trial to trial to create a high level of proactive interference (where interference from words on previous trials confuses retrieval for the current trial). They suggested that proactive interference impairs recall from long-term memory, but not the short-term store, thereby providing a clearer picture of what is going on in the short-term store. In contrast, in our Experiment 1 we sampled categories from a larger set and also sampled words within each category from a larger set. There was therefore less repetition of categories and of words across trials, thus arguably reducing the amount of proactive interference. With the additional assumption that primacy items are encoded more strongly into long-term memory because they have more time to be encoded, proactive interference could explain why we obtained more primacy than Davelaar et al.: Memory for the primacy part of the list could have been impaired by the higher level of proactive interference in Davelaar et al.

Experiment 2 used a similar method to Experiment 1, except that the same categories were used for each block of four consecutive trials so that proactive interference can build up within each block. The number of items to be remembered was reduced to 5 to further minimise any contribution from long-term memory. Probe type was postcued in this experiment for two reasons: First, postcueing rules out any effect of the kind of probe (category probe or spatial probe) on processes during encoding, ensuring that the serial-position effects are not confounded with strategic adjustments to the expected kind of probing. Second, Experiment 1 showed that, if anything, the precued method produces even more primacy than the postcued method. If we obtain primacy in tests of item memory with postcueing, we can be confident that this primacy is independent of whether or not people know in advance how they will be probed.

Method

Participants. Thirty participants, 15 women and 15 men between the ages of 17 and 35 years, took part in Experiment 2. They were all native speakers of English and participated to receive either course credit or £10 cash.

Design. Each trial included five English words from five different categories, presented visually in five different boxes on a computer screen. All five words were then probed for recall using either category probes or spatial position probes. There were 80 trials, grouped into 20 sets, with 4 trials in each set. The four trials in each particular set used words from the same five categories, so that proactive interference could build up within the set. No category was repeated on consecutive sets, so that proactive interference should be reduced between sets. Accordingly, we subdivided the trials in each set into one "first" trial, two "middle" trials, and one "last" trial per set. This meant that there were 20 "first" trials, 40 "middle" trials, and 20 "last" trials in total. Proactive interference should be minimal on the first trial and maximal on the last trial in each set.

Output order was varied randomly, with the constraint that for the "first" and "last" trials, every combination of input position and output position (e.g., word at input position 3 probed for recall at output position 2) occurred two times for each probe type, whereas for "middle" trials, every combination of input position and output position occurred four times for each probe type. This required 50 ($= 5 \times 5 \times 2$) words to be recalled, using 10 trials per probe type for "first" and "last" trials, and 100 ($= 5 \times 5 \times 4$) words, using 20 trials per probe type for "middle" trials. The spatial position of each word was determined at random on each trial.

The 20 sets of trials were constructed so that every combination of probe type for the "first" trial and probe type for the "last" trial ($2 \times 2 = 4$ combinations) occurred 5 times. Two "middle" trials were allocated randomly to each set, so that the probe type of the "last" trial in each set could not be predicted. After the trial structures and their order within each set had been fixed, categories and words were selected for them.

Sets of four trials were presented in random order in four test blocks of five sets (20 trials) each. The test blocks were preceded by eight practice trials, which were excluded from the data analysis.

Materials and procedure. The materials were the same as those in Experiment 1 except that only four words were used per category to increase proactive interference through increasing the likelihood of choosing the same word on different trials within a set. The timing intervals and procedure were identical to those of Experiment 1. Participants were tested for a 90-minute session consisting of four blocks, and they were encouraged to have a break after each block.

TABLE 4
Results of ANOVA for Experiment 2, testing effects of trial, probe type, input position, and output position

Variables	*Contrast*	F	MSE	p	η_p^2
Trial		7.339*	0.135	.011	.202
Probe type		218.484***	0.39	<.001	.883
Input	L	19.072***	0.364	<.001	.397
	Q	64.504***	0.05	<.001	.69
Output	L	154.502***	0.109	<.001	.842
	Q	36.927***	0.082	<.001	.56
Trial × Probe Type		2.076	0.14	.16	.067
Probe Type × Input	L × L	0	0.118	1	0
	L × Q	22.164***	0.043	<.001	.433
Probe Type × Output	L × L	6.136*	0.118	.019	.175
	L × Q	11.919**	0.059	.002	.291
Output × Input	L × L	4.032	0.101	.054	.122
	L × Q	2.201	0.097	.149	.071
	Q × L	18.376***	0.083	<.001	.388
	Q × Q	1.239	0.108	.275	.041
Trial × Probe Type × Input	L × L × L	0.451	0.062	.507	.015
	L × L × Q	1.217	0.047	.279	.04
Trial × Probe Type × Output	L × L × L	1.295	0.062	.264	.043
	L × L × Q	1.99	0.044	.169	.064
Trial × Output × Input	L × L × L	1.528	0.056	.226	.05
	L × L × Q	0.205	0.057	.654	.007
	L × Q × L	1.579	0.073	.219	.052
	L × Q × Q	0.008	0.089	.931	0
Probe Type × Output × Input	L × L × L	4.945*	0.042	.034	.146
	L × L × Q	4.722*	0.073	.038	.14
	L × Q × L	1.924	0.097	.176	.062
	L × Q × Q	0.704	0.117	.408	.024
Trial × Probe Type × Output × Input	L × L × L × L	1.164	0.048	.29	.039
	L × L × L × Q	0.521	0.046	.476	.018
	L × L × Q × L	0.464	0.087	.501	.016
	L × L × Q × Q	1.887	0.11	.18	.061

Note. The degrees of freedom in all cases are (1, 29). ANOVA = analysis of variance. η_p^2 = partial eta squared; L = linear contrast; Q = quadratic contrast; L × Q = interaction of linear contrast and quadratic contrast; Input = input position; Output = output position.
$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$.

Results

A factorial analysis of variance (ANOVA) was carried out with four within-subjects factors: trial position (first versus last in a set of four), probe type, input position and output position, and proportion of answers correct in the appropriate position as the dependent variable. The results from this ANOVA up to quadratic contrasts are presented in Table 4.

Figure 3 shows the input serial-position curve with extended primacy and modest recency and also the main effect of trial position (first versus last in a set), which did not interact with input position. These data show that there was a small effect of proactive interference (and release from proactive interference between sets) that affected all list positions approximately equally. Trial position did not interact with probe type, $F = 2.076$, $MSE = 0.14$, $p = .16$. Nevertheless, we tested the effect of trial position for each probe type separately. The effect was significant only for category probes, $F = 13.785$, $MSE = 0.085$, $p = .001$, but not for spatial probes, $F = 0.55$, $MSE = 0.189$, $p = .464$.

There was an interaction of probe type and input position, with greater primacy and recency for positional probes than category probes. The experiment also replicated the main effects of probe type (better recall for category probes) and the linear decline of recall over output position.

Discussion

There was an effect of proactive interference, but no evidence that it selectively suppressed

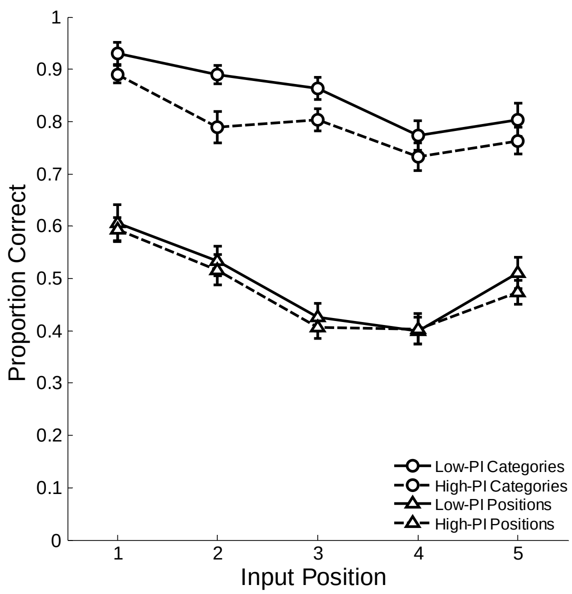

Figure 3. Accuracy as a function of probe type, proactive interference (low vs. high), and input position in Experiment 2, averaged over all output positions. Error bars represent within-subjects standard errors (e.g., Bakeman & McArthur, 1996).

performance in the primacy portion of the input position curve. The present data show that the primacy effect over input position is robust against a manipulation of proactive interference, for both relation memory and item memory.

This result rules out one more potential explanation for why we obtained a primacy effect in our test of item memory whereas Davelaar et al. (2005) found none, using essentially the same procedure. We could not think of any other difference between the experiment of Davelaar et al. and ours that might be held responsible for the different results. One might speculate that the fact that only a single item was tested, or that there was a time limit on recall (1.5 s) in the experiment of Davelaar et al., contributed to eliminating primacy, but there is no reason to believe that these features of the experiment have any impact on the primacy effect. We also need to consider the possibility that the experiment of Davelaar et al. simply had insufficient power to detect the primacy effect. As our analysis of the first-probed item in Experiment 1 has shown, the primacy effect for the first-recalled item is relatively small, so that it could be missed due to a lack of power (Davelaar et al., 2005, had only 30 trials in the relevant presentation-rate condition, compared to our 56 in Experiment 1). We cannot conclusively resolve the remaining discrepancy between our results and those of Davelaar et al. Fortunately, this discrepancy is of little consequence for the question of what causes serial-position effects in list memory. As we explain below, our investigation of various potential sources of serial-position effects does not hinge on there being both primacy and recency effects on item memory: As long as at least one of these effects is obtained, our conclusions hold.

To conclude, we obtained consistent evidence across two experiments that there is both primacy and recency for item memory as well as relation memory. This result was also obtained with two further experiments of ours, not reported here, using the same paradigm. Our result is in agreement with previous studies testing item memory through different methods (e.g., Cowan et al., 2002; Hay et al., 2007; McElree & Dosher, 1989; Monsell, 1978; Oberauer, 2003). Therefore, we are confident that there are both primacy and recency effects over input position in tests of item memory as well as relation memory. We now consider possible explanations of these effects, undertaking a systematic exploration of potential mechanisms and retrieval processes through computational modelling.

MODELLING CAUSES OF SERIAL-POSITION EFFECTS

We constructed a generic computational framework to compare the different potential loci of serial-position effects (for the Matlab code, see Figure S1, which is available via the supplementary tab on the article's online page at http://dx.doi.org/10.1080/09658211.2012.726629). We did not aim to quantitatively fit the data. Rather, we investigated whether the qualitative pattern of our data—in particular the U-shaped serial-position curve over input for both item and relation memory—can be reproduced by implementing serial-position effects at different loci in the model (i.e., strength of item representations, strength of bindings between items and input positions, strength of bindings between items and spatial locations, distinctiveness of input positions, and distinctiveness of spatial locations). Any combinations of loci and retrieval modes that cannot generate these U-shaped serial-position curves for both kinds of probe are ruled out by the present data.

All hypothetical causes of serial-position effects explored in our simulations were assumed to vary across serial position according to a U-shaped function. For instance, simulations investigating whether variation in item activation contributes to serial-position effects assumed that items at the beginning and the end were more strongly activated than those in the middle of the list. This was intended to give every mechanism the best chance of producing a U-shaped serial-position curve, so that any mechanisms that could not, even under these generous conditions, could be discounted. We evaluated all simulations for whether they produced U-shaped serial-position curves over input position.

We used a generic architecture in which serial order is represented by associating item representations to context markers; this assumption is shared by the most successful models of serial recall (Lewandowsky & Farrell, 2008). Within this architecture, we considered six sources of serial-position effects in conjunction with four retrieval mechanisms (see Table 1). They are instantiated in various combinations in published models.

Serial-position effects in item memory

Figure 4 demonstrates the major structures involved in our model of memory for items. There are item representations, denoted by letters; position markers representing ordinal positions in the input sequence, denoted by numbers; and bindings between the two. The position markers can overlap each other (the shaded areas on Figure 4). Therefore activation of Position Marker 1 will also partially activate Position Marker 2 to the extent that they overlap. The amount of activation that feeds from each position marker to the item presented in that position depends on the strength of binding between the position and item. The relative amount of activation that feeds to neighbouring items depends on positional distinctiveness, as modelled by the proportion of overlap between neighbouring position markers.

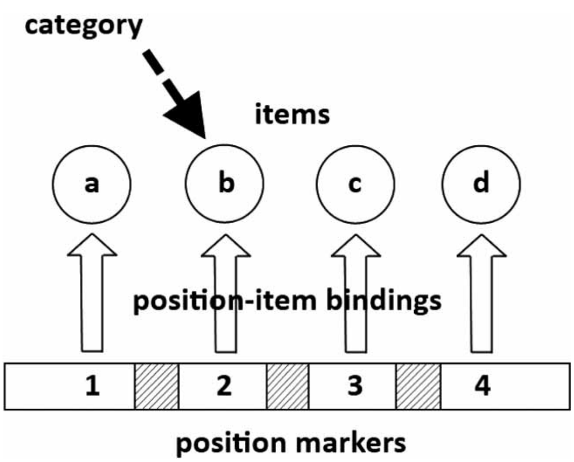

Figure 4. Representations involved in the model of memory for items: Overlapping position markers denoted by numbers, item representations denoted by letters, and bindings between the two. A category cue is depicted giving extra activation to Item b.

During retrieval, the category cue used in the experiment was assumed to provide extra activation to the appropriate list item, but not to any other list items (see Figure 4). Activation had to surpass a threshold for retrieval to occur. This threshold was set higher than the activation provided by the category cue alone, reflecting the aim to recall an item that is both from the correct category and from the current list. Gaussian noise in the model contributed to whether item activation exceeded the threshold or not. The added noise could occasionally lift an extra-list item matching the currently cued category above threshold, thereby producing an error (i.e., an extralist intrusion), but these were not explicitly modelled.

Two retrieval mechanisms were considered in the model. The first makes no use of the item–position bindings. It adds the activation from the cued category (which is positive only for the one matching item) to the intrinsic activations of the items themselves (i.e., the activation they received at encoding) and recalls the item with the highest resulting activation, provided it surpasses the threshold. The second mechanism implements retrieval by scanning through the list. It uses the position markers one by one in forward order to cue items. At each step, activation conferred to the items from the current position cue is added to the intrinsic item activation and the activation from the category cue. When one or more items exceed the threshold, the item with the highest resulting activation is retrieved, and the scanning procedure finishes. If no item exceeds the threshold, the next position marker is activated, and so on until the end of the list. An omission is returned if no items are recalled after all position markers have been used as cues (i.e., scanning has reached the end of the list). We combined these two retrieval processes with three possible locations of serial-position effects: memory strength (i.e., activation) of items, strength of bindings between items and positions, and distinctiveness of positions.

The findings from simulations with the model were clear: primacy and recency gradients on the intrinsic activation of item representations produced a U-shaped serial-position curve with either retrieval mechanism. Primacy and recency over the strength of item–position bindings produced a U-shaped serial-position curve only with retrieval by scanning. Variation in positional distinctiveness (i.e., higher overlap of medium than of end positions) only mattered with retrieval by scanning, and it could *not* produce a U-shaped curve. In fact, higher positional distinctiveness at the ends than in the middle of the list led to better recall of items at *medial* positions. This is because overlap in the position markers gives items extra chances of being recalled (when cued with nearby position markers). In tests of item memory, it does not matter if an item is recalled in response to the ‘‘wrong’’ position marker as cue, it only matters that the item is recalled at all. Medial position markers have more overlaps with other markers than those at the extreme ends, and therefore medial items have more opportunities of being activated above threshold, leading to an inverted-U serial-position curve.

Serial-position effects in relation memory

Figure 5 shows the structures involved in the model of relation memory. It is an extension of the model for item memory that includes representations of spatial locations. The shaded areas represent overlap in memory for nearby locations. In addition to their bindings to input-position markers, items were also bound to spatial locations. Moreover, spatial locations were bound to input-position markers, such that the spatial location in which the first-presented item appeared was bound to the first position marker, and so on. In Figure 5, these bindings are only shown for the first serial position to avoid clutter, but there were such bindings for all serial positions. As in the experiments described above, spatial location and input position were uncorrelated across trials, and therefore we assigned spatial locations to input positions at random in the simulations. In addition to the three sources of serial-position effects discussed in the context of item memory (i.e., intrinsic activation of items, bindings between items and input positions, and distinctiveness of input positions), serial-position effects in relation memory could be caused by variation in the *strength of item–location bindings*, the *strength of bindings between spatial locations and input positions*, and by variation in *spatial distinctiveness* (i.e., by varying how much adjacent locations overlap in memory, as for the input positions above).

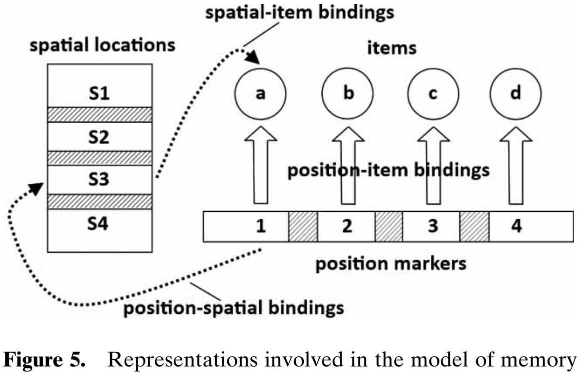

Figure 5. Representations involved in the model of memory for relations: an extension of the model for items adding spatial locations, denoted by ‘‘S’’ numbers, spatial–item bindings, and input position–spatial bindings. Only one example of spatial–item and position–spatial bindings is shown to avoid clutter, where in the model there would be four of each.

We explored two retrieval mechanisms. The first is to use the probed spatial location as a cue to directly retrieve the item linked to it through item–location bindings. This worked in the same way as cueing with input position above. The cued spatial location partially activated any overlapping locations, and activation was forwarded from locations to items according to the strength of the item–location bindings. After the addition of the intrinsic strength of item representations and Gaussian noise, the most activated item would be retrieved as long as its activation exceeded a threshold.

The second retrieval procedure uses scanning through the list in order of presentation. People could scan through the input-position markers in forward order, simultaneously recalling the spatial location and the item that is bound to each position marker until they find the location that matches the given location probe (Cowan et al., 2006). The item retrieved at the same time would then be reported. Such an indirect retrieval process is conceivable if retrieval in order of presentation is much easier than retrieval of an item by its direct binding to the spatial cue. There is some evidence that retrieval in forward order is easier than retrieval in any other order

(Lange, Cerella, & Verhaeghen, 2011; Lange, Verhaeghen, & Cerella, 2010), rendering retrieval by scanning plausible. Scanning involved stepping through the input position markers one at a time and retrieving the most activated spatial location. Activation was conferred to spatial locations from the position markers through the bindings between them; the degree of activation for each location depended on the overlap between neighbouring input positions and the strength of bindings between input positions and spatial locations. Once a location was retrieved that matched the probed location, the current input-position marker was used to cue for its associated item as described in the previous section. If any item's activation exceeded the threshold at this step, the item with the highest activation was selected as output, otherwise an omission was returned.

Again, primacy and recency in the intrinsic activation of item representations produced a U-shaped serial-position curve with either retrieval mechanism. With direct cueing using a spatial location, only the strength of bindings between items and spatial locations and the distinctiveness of spatial locations could have a potential impact on the serial-position curve over input position. In fact, spatial distinctiveness had no effect on the serial-position curve over input order because spatial locations were uncorrelated with input positions. Strength of bindings between items and spatial locations affected primacy and recency only if the bindings were stronger at more extreme input positions; it did not matter if the bindings were stronger at more extreme spatial locations.

Retrieval by scanning through input positions tended to produce a slightly asymmetrical serial-position curve, weighted towards primacy. This was due to the forward-scanning nature of the mechanism. For example, assume that a participant is cued with the fourth spatial location, which in this trial corresponded to the final input position. Due to noise in the retrieval process, the fourth spatial location might be retrieved already by an earlier input position marker. Here the scanning stops, and an attempt is made to retrieve the item at that input position. The mechanism is blind to the fact that a stronger activation of the fourth spatial location would probably have occurred had it been cued with the final input position. Earlier input positions suffer less from this problem as they have a greater opportunity to cue their most strongly related spatial location.

With scanning as the retrieval mechanism, primacy and recency in the strength of bindings between items and spatial locations, or the strength of bindings between items and positions, produced a U-shaped serial-position curve. Any overlap (or reduced distinctiveness) in the input position markers also produced a U-shaped curve. This is because recalling items in the correct positions is important for relation memory. Medial positions have more overlaps, leading to a higher chance of recalling the wrong spatial location, or the wrong item, either of which results in a mismatch of probed location and recalled item.

Table 1 summarises the results of our simulations. The combinations of retrieval processes and presumed mechanisms of serial-position effects that are viable in light of our data are printed in bold. The table shows that our findings rule out about half of the theoretical possibilities.

GENERAL DISCUSSION

The goal of the present experiments was to analyse serial-position effects in short-term recall separately for item memory and for relation memory. The experimental design allowed a direct comparison of tests of item memory (using category probes) with tests that unambiguously required relation memory (binding of items to their spatial position at presentation) in a within-subjects comparison that avoids confounds with other variables, and using a procedure that separates input from output order. Based on previous literature using category probes (Davelaar et al., 2005) or item recognition (Monsell, 1978; Oberauer, 2003), we expected item memory to show mostly recency and little primacy. Contrary to this expectation, our results with both probe types showed primacy effects that were at least as strong as corresponding recency effects.

Serial-position effects over input position

Our design provided a separation of serial-position effects along the dimensions of input and output position. This allows us to investigate the effects over input position controlled for the impact of output position. The separation of item memory and relation memory helps us to further narrow down the set of candidate mechanisms for serial-position effects.

As demonstrated by our simulations, primacy and recency in item memory cannot be caused by a mechanism involving positional confusions/ distinctiveness alone. Such a mechanism predicts an inverted-U shaped serial-position curve, contrary to the U-shaped curve obtained in the data. Distinctiveness is a crucial component of several current memory models—for example, the scale-independent memory, perception, and learning model (SIMPLE; Brown et al., 2007), the start–end model (Henson, 1998), and the model of Botvinick and Watanabe (2007). To account for a U-shaped serial-position curve in item memory, distinctiveness models are required to make additional assumptions. For SIMPLE (Brown et al., 2007) a threshold for distinctiveness is proposed. Items whose positions are more confusable with each other (e.g., medial items in a list) might fall below the threshold and therefore not be recalled at all. This seems to us an implausible assumption because it implies an inefficient use of memory. It would be more helpful to attempt to remember these items, but potentially confuse them and recall them in the incorrect positions. As position does not matter for item memory, this would lead to more chances to correctly recall medial items, resulting in an inverse-U serial-position curve, as shown in our model.

In the start–end model, in addition to differential distinctiveness, "strength" is introduced into the calculation leading to the selection of an item for recall. Extreme position markers have greater strength than more medial markers. Items that fall below a threshold of strength are omitted. This strength parameter in the start–end model is akin to strength of bindings between items and input positions in our model. Without it, the start–end model would be unable to produce a U-shaped serial-position curve for item memory.

Explanations of primacy and recency effects over input positions based on variations in the strength of bindings between items and input position markers face an obvious difficulty: Neither successful item recall nor successful recall of item–location relations requires information about an item's input position. If we assume that retrieval is accomplished through direct bindings between the relevant cues (i.e., the category or the spatial location) and the target item, then our test of item memory would be sensitive only to differences in the item's intrinsic activation; our test of relation memory would in addition be sensitive to variations in the strength of bindings between items and spatial locations. The strength of bindings between items and their input position markers can play a role only if we assume a more indirect retrieval process based on scanning through the list in its order of presentation.

Our experiments show largely parallel serial-position curves for tests of item memory and order memory. Parsimony therefore strongly suggests that the same mechanisms are responsible for serial-position effects in both kinds of tests. Table 1 shows that only two models are compatible with the U-shaped serial-position curve for both item and relation memory. One is a model with serial-position effects on the intrinsic activation of items. The other is a model with serial-position effects on the strength of bindings between items and their input positions, combined with retrieval by scanning.

The implications of our modelling results do not depend on the shape of the serial-position curve; in particular they do not depend on there being both primacy and recency effects in the data. We implemented symmetric U-shaped serial-position gradients on different parameters in the models (because this is the shape we observed in the experiments), and as a consequence we obtained either U-shaped serial-position curves, or serial-position curves that did not reflect the implemented serial-position gradients at all (i.e., they were either flat or had an inverted U shape). Therefore, our simulations show that serial-position gradients on certain model parameters (in particular, item activation and bindings between items and input positions) translate into corresponding serial-position effects in the data, whereas serial-position gradients on other parameters do not. This result is independent of the shape of the serial-position gradients. Therefore, it does not matter whether there is both primacy and recency in item memory and relation memory—as long as there is either primacy or recency (or both) in item memory as well as in relation memory, our conclusions hold.

To conclude, our data place constraints on explanations of serial-position effects in immediate memory. The U-shaped serial-position curve over input position for both item and relation memory cannot be explained purely by positional distinctiveness. Either these serial-position effects arise from variations in the intrinsic strength (or activation) of the items, or they arise from variations in the strength of bindings between items and input positions. In the latter case, memory must rely on retrieval by scanning through the list in the order of presentation, rather than on direct bindings between items and the given retrieval cues.

Manuscript received 1 March 2012
Manuscript accepted 29 August 2012
First published online 24 September 2012

REFERENCES

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*, *47*, 207–231.

Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1995). The CELEX lexical database (Release 2) [CD ROM]. Philadelphia, PA: University of Pennsylvania, Linguistic Data Consortium.

Bakeman, R., & McArthur, D. (1996). Picturing repeated measures: Comments on Loftus, Morrison, and others. *Behavior Research Methods, Instruments, & Computers*, *28*, 584–589.

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.

Bjork, E. L., & Healy, A. F. (1974). Short term order and item retention. *Journal of Verbal Learning and Verbal Behaviour*, *13*, 80–97.

Botvinick, M., & Watanabe, T. (2007). From numerosity to ordinal rank: A gain-field model of serial order representation in cortical working memory. *Journal of Neuroscience*, *27*, 8636–8642.

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.

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: A network model of the phonological loop and its timing. *Psychological Review*, *106*, 551–581.

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.

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

Cowan, N., Saults, J. S., & Morey, C. C. (2006). Development of working memory for verbal–spatial associations. *Journal of Memory and Language*, *55*, 274–289.

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

Duncan, M., & Murdock, B. (2000). Recognition and recall with precuing and postcuing. *Journal of Memory and Language*, *42*, 301–313.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. A. Bussenius, Trans.). New York, NY: Dover Publications. (Original work published 1885)

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

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

Farrell, S., & Lelièvre, A. (2009). Is scanning in probed order recall articulatory? *Quarterly Journal of Experimental Psychology*, *62*, 1843–1858.

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

Fuchs, A. H. (1969). Recall for order and content of serial word lists in short-term memory. *Journal of Experimental Psychology*, *82*, 14–21.

Grossberg, S., & Stone, G. (1986). Neuronal dynamics of attention switching and temporal order information in short term memory. *Memory & Cognition*, *14*, 451–468.

Hay, D. C., Smyth, M. M., Hitch, G. J., & Norton, N. J. (2007). Serial position effects in short-term visual memory: A SIMPLE explanation? *Memory & Cognition*, *35*, 176–190.

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

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

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

Lange, E. B., Cerella, J., & Verhaeghen, P. (2011). Ease of access to list items in short-term memory depends on the order of the recognition probes. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *37*, 608–620.

Lange, E. B., Verhaeghen, P., & Cerella, J. (2010). Dual representation of item positions in verbal short-term memory: Evidence for two access modes. *European Journal of Cognitive Psychology*, *22*, 463–479.

Lee, C. L., & Estes, W. K. (1977). Order and position in primary memory for letter strings. *Journal of Verbal Learning and Verbal Behaviour*, *16*, 395–418.

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

Li, S.-C., & Lewandowsky, S. (1993). Intralist distractors and recall direction: Constraints on models of memory for serial order. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *19*, 895–918.

Li, S.-C., & Lewandowsky, S. (1995). Forward and backward recall: Different retrieval processes. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *21*, 837–847.

Madigan, S. A. (1971). Modality and recall order interactions in short-term memory for serial order. *Journal of Experimental Psychology*, *87*, 294–296.

McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term-memory—The time course of recognition. *Journal of Experimental Psychology: General*, *118*, 346–373.

Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*, *10*, 465–501.

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

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

Nipher, F. E. (1876). Probability of error in writing a series of numbers. *American Journal of Science*, *12*, 79–80.

Nipher, F. E. (1878). On the distribution of errors in numbers written from memory. *Transactions of the Academy of Science of St. Louis*, *3*, ccx–ccxi.

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

Oberauer, K., & Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, *55*, 601–626.

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

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

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

Van Overshelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An updated and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language*, *50*, 289–335.