## How Is the Serial Order of a Spatial Sequence Represented? Insights From Transposition Latencies

Mark J. Hurlstone and Graham J. Hitch

CITATION

Hurlstone, M. J., & Hitch, G. J. (2014, December 1). How Is the Serial Order of a Spatial Sequence Represented? Insights From Transposition Latencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. Advance online publication. http://dx.doi.org/10.1037/a0038223

# How Is the Serial Order of a Spatial Sequence Represented? Insights From Transposition Latencies

Mark J. Hurlstone
University of York and University of Western Australia

Graham J. Hitch
University of York

How is the serial order of a spatial sequence represented in short-term memory (STM)? Previous research by Farrell and Lewandowsky (Farrell & Lewandowsky, 2004; Lewandowsky & Farrell, 2008) has shown that 5 alternative mechanisms for the representation of serial order can be distinguished on the basis of their predictions concerning the response times accompanying transposition errors. We report 3 experiments involving the output-timed serial recall of sequences of seen spatial locations that tested these predictions. The results of all 3 experiments revealed that transposition latencies are a negative function of transposition displacement, but with a reduction in the slope of the function for postponement, compared with anticipation errors. This empirical pattern is consistent with that observed in serial recall of verbal sequences reported by Farrell and Lewandowsky (2004), and with the predictions of a competitive queuing mechanism, within which serial order is represented via a primacy gradient of activations over items combined with associations between items and positional markers, and with suppression of items following recall. The results provide the first clear evidence that spatial and verbal STM rely on some common mechanisms and principles for the representation of serial order.

*Keywords:* competitive queuing, serial order, short-term memory, spatial, transposition latencies

A pivotal issue to be addressed by any adequate theory of short-term memory (STM) is the problem of serial order (Lashley, 1951): how people store and reproduce the order of a sequence of items. From vocabulary acquisition (Baddeley, Gathercole, & Papagno, 1998) to the learning and production of action sequences (Agam, Bullock, & Sekuler, 2005; Agam, Galperin, Gold, & Sekuler, 2007; Baddeley, 2007), this competency is thought to be crucial for numerous acts of higher level cognition. According to one popular theory—namely, the working memory model of Baddeley and Hitch (1974)—STM for a sequence of verbal items depends upon the action of a phonological loop—a rapidly decaying phonological store, the contents of which can be revivified via a compensatory articulatory control process (Baddeley, 1986)—whereas STM for a spatial sequence depends upon the action of a visuospatial sketchpad—a rapidly decaying visual store, the contents of which can be revivified via a compensatory spatial rehearsal process (Logie, 1995). Although the working memory model has proved successful in explaining a wealth of data obtained with serial tasks, a major shortcoming of the model is that it does not specify how serial order information is represented within the two STM systems.

In response to these shortcomings of the conceptual model, several authors have developed computational theories of the phonological loop that quantitatively instantiate its core assumptions, while specifying explicit mechanisms for serial order (Burgess & Hitch, 1999; Henson, 1998a; Page & Norris, 1998). The successes and failures of these theories have inspired other authors to develop computational theories of verbal STM for serial order, albeit cast outside the phonological-loop/working-memory framework (Botvinick & Plaut, 2006; Brown, Neath, & Chater, 2007; Brown, Preece, & Hulme, 2000; Farrell & Lewandowsky, 2002; Lewandowsky & Farrell, 2008). The seriating mechanisms that should feature within an adequate computational account of the phonological loop have been discovered via comparisons between these models on the basis of their core constructs (Farrell & Lelièvre, 2009; Farrell & Lewandowsky, 2004, 2012; Henson, 1998b, 1999; Hitch, Fastame, & Flude, 2005; Lewandowsky & Farrell, 2008; Oberauer, 2003). In contrast, computational theories of STM for spatial sequences have yet to be developed, and the principles and mechanisms governing the representation of serial order in spatial STM have remained elusive. This is perhaps surprising, given that very many lines of evidence now indicate that spatial and verbal STM for serial order are functionally similar (see Hurlstone, Hitch, & Baddeley, 2014, and Parmentier, 2010, for reviews), suggesting that principles of serial order in verbal STM may generalize to the spatial domain.

This article examines the representation of serial order in spatial STM using a combination of experimentation and modeling. Specifically, it examines whether principles and mechanisms previ-

ously shown to contribute to the representation of serial order in verbal STM also contribute to the representation of serial order in spatial STM. We employ a generic modeling architecture to contrast the predictions of five computational models of serial order built from different combinations of representational principles widely employed in models of verbal STM. Although the models cannot be distinguished on the basis of their predicted response probabilities, they make interestingly different predictions concerning the response times accompanying transposition errors. We report three experiments involving the recall of sequences of spatial locations that tested the predictions of the models. To foreshadow, the results consistently support the predictions of a representational mechanism embodying a primacy gradient, position marking, and response suppression. This mechanism has previously been identified in a kindred study involving the recall of verbal sequences (Farrell & Lewandowsky, 2004), suggesting that common principles underpin serial order across the spatial and verbal domains. We discuss the broader theoretical implications of this outcome for the representation of serial order within working memory.

## Similarities Between Serial Order in the Spatial and Verbal Domains

Memory for serial order is typically examined with the serial recall task, in which participants are given a sequence of items that they must subsequently recall in the correct order. In the verbal domain—where the task has been most frequently employed—the stimuli typically consist of letters, digits, or words, whereas in the spatial domain, the stimuli typically consist of spatial locations or spatial movements. The measure of spatial serial recall that has been most commonly utilized is the Corsi-Blocks Task (Corsi, 1972). In this task, an experimenter taps blocks arranged irregularly on a wooden tablet in a sequence that an observing participant must subsequently reproduce. In computerized versions of the task, the locations are represented by two-dimensional squares and a sequence is conveyed by briefly highlighting each location in turn. The participant is then required to reproduce the sequence by pointing to the locations using a touch-sensitive display or a mouse-driven pointer. A major advantage of computerized versions of the task is that they permit the recording of response latencies (De Lillo, 2004; Fischer, 2001). One objective of many studies employing this task—and basic variants of it (viz., the dots task; Jones, Farrand, Stuart, & Morris, 1995)—has been to establish whether serial order is processed in similar ways across the spatial and verbal domains. The strategy adopted in pursuit of this goal has been to determine whether major phenomena of serial order witnessed in verbal STM are extensible to spatial STM. The two phenomena that have received the most empirical scrutiny are serial position curves and transposition gradients.

The serial position curve plots recall accuracy as a function of the serial position of items. When plotted for verbal sequences, the forward recall curve exhibits two canonical features. First, there is a sharp monotonic decrease in accuracy extending from the first position onward—known as the primacy effect. Second, there is an upturn in the trend line at the final serial position—known as the recency effect. Serial position curves can also be plotted using interresponse time as the dependent measure. Response timing studies have shown that people leave a long pause prior to outputting the first item in a verbal sequence, after which recall latency varies inversely with serial

position accuracy giving rise to an inverted U-shaped latency serial position curve (Anderson, Bothell, Lebiere, & Matessa, 1998; Farrell & Lewandowsky, 2004; Maybery, Parmentier, & Jones, 2002; Parmentier & Maybery, 2008; Thomas, Milner, & Haberlandt, 2003).

The forward recall curve for verbal sequences is associated with a highly systematic pattern of errors. The majority of errors are transpositions (Aaronson, 1968; Henson, 1996), which occur when items are recalled in the wrong positions. Transpositions can be classified according to their displacement—the numerical difference between an item's input and output positions. Transpositions with negative displacement values are known as "anticipations" and occur when an item is recalled before its target position. For example, a $-2$ displacement indicates an item has been recalled two positions ahead of its correct position. Transpositions with positive displacement values are known as "postponements" and occur when an item is recalled after its target position. For example, a $+4$ displacement indicates an item has been recalled four positions after its correct position. Transposition gradients plot the probability of transpositions as a function of their displacement and exhibit three hallmark characteristics (Farrell & Lewandowsky, 2004). First, the gradients peak at displacement 0, indicating that the majority of responses are correct. Second, the probability of a transposition decreases as the absolute displacement increases. Thus, when an item is recalled in the wrong position it will tend to be close to its correct position—the locality constraint (Henson, 1996; Henson, Norris, Page, & Baddeley, 1996). Third, the error gradients for anticipations and postponements are approximately symmetrical (Farrell & Lewandowsky, 2004): The probability of a transposition at each absolute displacement is similar for anticipations and postponements.

These features of memory for serial order are not confined to the verbal domain: accuracy serial position curves (Avons, 2007; Farrand, Parmentier, & Jones, 2001; Guérard & Tremblay, 2008; Jones et al., 1995; Tremblay, Guérard, Parmentier, Nicholls, & Jones, 2006), latency serial position curves (Parmentier, Andrés, Elford, & Jones, 2006; Parmentier, Elford, & Maybery, 2005), and transposition gradients (Hurlstone, 2010; Jalbert, Saint-Aubin, & Tremblay, 2008; Parmentier et al., 2006; Smyth & Scholey, 1996) exhibiting the empirical features just reviewed have also been observed for serial recall of sequences of spatial locations. Indeed, spatial and verbal serial recall share many other functional similarities, including similar distributions of item and order errors (Guérard & Tremblay, 2008) and similar effects of sequence length (Jones et al., 1995; Smyth, 1996; Smyth & Scholey, 1994, 1996), temporal grouping (Parmentier et al., 2006), item similarity (Jalbert et al., 2008), and sequence repetition (Couture & Tremblay, 2006). These similarities between spatial and verbal STM are suggestive of the operation of common principles for representing serial order in the spatial and verbal domains. In the next section, we outline mechanisms and principles that have been implicated in the representation of serial order in verbal STM, which may also be implicated in the representation of serial order in spatial STM.

## Mechanisms and Principles for Serial Order

There has been some theoretical convergence between the various computational models of verbal STM, and several mechanisms and principles for serial order can be discerned. Table 1 lists those theories and classifies them according to the core mechanisms and principles they instantiate. It can seen from inspection of this table that most

Table 1
*Contemporary Theories of Verbal Short-Term Memory for Serial Order and the Mechanisms and Representational Principles They Instantiate*

| Model | Competitive queuing | Position marking | Primacy gradient | Response suppression | Output interference |
|---|---|---|---|---|---|
| SRN (Botvinick & Plaut, 2006) | ✓ | X | X | X | X |
| SIMPLE (Brown et al., 2007) | X | ✓ | X | X | ✓ |
| OSCAR (Brown et al., 2000) | ✓ | ✓ | ✓ | ✓ | ✓ |
| Burgess & Hitch (1999) | ✓ | ✓ | ✓ | ✓ | X |
| SEM (Henson, 1998a) | ✓ | ✓ | ✓ | ✓ | X |
| SOB (Farrell & Lewandowsky, 2002) | X | X | ✓ | ✓ | X |
| C-SOB (Lewandowsky & Farrell, 2008) | X | ✓ | ✓ | ✓ | ✓ |
| LIST PARSE (Grossberg & Pearson, 2008) | ✓ | X | ✓ | ✓ | X |
| Primacy model (Page & Norris, 1998) | ✓ | X | ✓ | ✓ | X |

*Note.* SRN = simple recurrent network; SIMPLE = scale-invariant memory, perception, and learning; OSCAR = oscillator-based associative recall; SEM = start-end model; SOB = serial-order-in-a-box; C-SOB = context-serial-order-in-a-box; LIST PARSE = laminar integrated storage of temporal patterns for associative retrieval, sequencing, and execution.

theories generate serial order using a parallel sequence planning and control mechanism known as competitive queuing. However, the theories differ in terms of the exact principles they use to represent serial order within the competitive queuing system, with some representing serial order by associating items to an index of their sequence position using position marking; some by incorporating a primacy gradient of activations over items; and others by a combination of position marking and a primacy gradient. Additionally, almost all theories implement response suppression, and some postulate a role for output interference. We turn now to a description of the different mechanisms and principles.[1]

## Competitive Queuing

A schematic of a generic competitive queuing mechanism (e.g., Bullock, 2004; Bullock & Rhodes, 2003; Davelaar, 2007; Houghton, 1990) envisaged as a neural network model is illustrated in Figure 1. The model consists of two layers of localist item nodes: a parallel planning layer and a competitive choice layer. The nodes in the parallel planning layer represent the pool of elements from which sequences are composed. Recalling a sequence is a two-stage process. In the first stage, an ordering mechanism activates in parallel a subset of the nodes in the planning layer, with the relative strength of node activations coding the relative output priority of items. In the second stage, these activations are projected to corresponding nodes in the competitive choice layer. The node activations in this layer are governed by recurrent-competitive-field dynamics, such that each item node excites itself and sends lateral inhibition to competitor nodes in the same layer. This sets up a response competition, and the item with the strongest activation level is chosen for recall, after which a feedback signal from the competitive choice layer inhibits its corresponding representation in the parallel planning layer. This process iterates until recall of the entire sequence is completed.

There are several variations on the competitive queuing mechanism just described. The most fundamental concerns the mechanism employed by competitive queuing models to generate the activation gradient used to represent serial order in the parallel planning layer. In the simplest models, a static mechanism is employed to generate a single activation gradient (viz., a primacy gradient; see later), whereas in more sophisticated models, a dy-

namic mechanism is employed that modulates the activation gradient over time (viz., position marking; see next). Depending upon the type of representation of serial order utilized, models also differ in the extent of their reliance on the postretrieval inhibition of items (viz., response suppression; see later).

## Position Marking

Position marking is an approach to representing serial order in which sequence items are associated during encoding with some independent and dynamically varying representation of temporal (Brown et al., 2000, 2007; Burgess & Hitch, 1999), absolute (Burgess & Hitch, 1992; Lewandowsky & Farrell, 2008), or relative (Henson, 1998a) within-sequence position. The positional representations are only approximate, meaning that the representations of neighboring positions overlap to some degree. At recall, the positional cues are reinstated in turn and items are activated by an amount dependent upon the similarity between the current positional cue and the positional cue they were originally associated with.

## Primacy Gradient

A simpler scheme for representing serial order is in terms of a primacy gradient of activations. During serial order encoding, the first item in the sequence is activated strongest, with the activations of subsequent items decreasing monotonically across input positions. In some models (e.g., Farrell & Lewandowsky, 2002; Grossberg, 1978; Grossberg & Pearson, 2008; Page & Norris, 1998), serial order is represented on the basis of the primacy

---

[1] One representational principle that we do not consider here is associative chaining (Ebbinghaus, 1964; Lewandowsky & Murdock, 1989). Chaining is a representational scheme in which serial order is coded by forming associations between successive items in a sequence. Ordered recall then proceeds by traversing these associations, which act as the cues for sequence production. Although once a popular approach, there is now a wealth of evidence indicating that people do not use chained associations to represent serial order in STM (see Hurlstone, Hitch, & Baddeley, 2014, for a review; see also Solway, Murdock, & Kahana, 2012, for recent data purporting to support a role for chaining, and Farrell, Hurlstone, & Lewandowsky, 2013, for a reappraisal of that evidence).
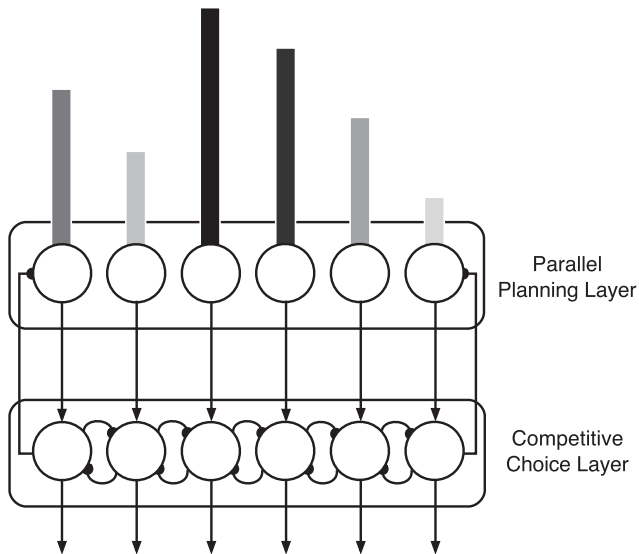
*Figure 1.*   Schematic of a two-layer competitive queuing sequence planning and control mechanism comprising a parallel planning layer (upper field of nodes) and a competitive choice layer (lower field of nodes). Lines terminating with arrows represent excitatory connections, whereas lines terminating with semicircles represent inhibitory connections. Note that each node in the lower competitive choice layer has an inhibitory connection to every other node in the same layer, but for simplicity, only adjacent-neighbor inhibitory connections are shown. Similarly, each node in the competitive choice layer has an inhibitory connection to its corresponding node in the parallel planning layer, but to avoid visual clutter, only feedback connections for the leftmost and rightmost nodes are illustrated. See the main text for further details.

gradient alone. When complemented by response suppression (see later), such models accomplish forward serial recall via an iterative process of selecting the most active item, before suppressing its activation, so that the next strongest item can be chosen. Other models incorporate a primacy gradient in conjunction with position marking to represent serial order (e.g., Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998a; Lewandowsky & Farrell, 2008), giving rise to a more robust sequencing mechanism.

## Response Suppression

Response suppression refers to the inhibition of items once they have been recalled and represents an almost universal assumption of theories of verbal STM (e.g., Brown et al., 2000; Burgess & Hitch, 1999; Farrell & Lewandowsky, 2002; Grossberg & Pearson, 2008; Henson, 1998a; Lewandowsky & Farrell, 2008; Page & Norris, 1998). In competitive queuing models, response suppression is represented by the inhibitory feedback signal from the competitive choice layer to the parallel planning layer following the recall of an item. Response suppression is a crucial ingredient in competitive queuing models that represent serial order using a primacy gradient, because, in its absence, the selection mechanism would perseverate on the initial response, which would always remain the most active. Response suppression is less crucial in competitive queuing models that represent serial order using position marking, because the re-

evolving positional context signal dynamically updates the activations in the parallel planning layer, thereby reducing the burden on response suppression for sequencing. Nevertheless, even models that represent serial order using position marking typically incorporate response suppression in order to minimize the occurrence of erroneous repetition errors, which would otherwise be abnormally frequent.

## Output Interference

Output interference is not a representation of serial order, but rather an ancillary assumption incorporated in some theories of STM to more accurately model primacy and sequence length effects in serial recall (Brown et al., 2000, 2007; Lewandowsky & Farrell, 2008). It refers to the notion that the recall of an item from STM interferes with the representations or accessibility of items that are yet to be retrieved. This output interference occurs irrespective of whether or not a recalled item is subsequently suppressed, and its effects accumulate as sequence production unfolds, meaning that the representations of items in the middle and toward the end of the sequence will be most impaired by its action.

## The Current Study

Given the empirical similarities between spatial and verbal STM for serial order, it seems likely that at least some of the principles and mechanisms just described also contribute to the representation of serial order in spatial STM. The question is, which ones, if not all of them? In a recent review of the literature, Hurlstone et al. (2014) proposed that all short-term memories (e.g., spatial, verbal, visual) utilize the competitive queuing mechanism to plan, represent, and recall sequences. This claim was founded on three precedents. First, competitive queuing models offer a natural account of the locality constraint governing transpositions in serial tasks. Thus, if the activations of elements in one or both layers of the competitive queuing output module are perturbed by moderate random noise, this will alter the relative output priority of items, and, because of the gradient-based representation of serial order, adjacent-neighbor movement errors will predominate. Second, the competitive queuing mechanism has received direct support from electrophysiological recording data obtained with monkeys engaged in a spatial imitation task (Averbeck, Chafee, Crowe, & Georgopoulos, 2002, 2003; Averbeck, Crowe, Chafee, & Georgopoulos, 2003), and aspects of these data have been corroborated in ERP studies of humans (Agam, Huang, & Sekuler, 2010; Agam & Sekuler, 2007). Third, competitive queuing models have been developed and successfully applied to a variety of serial performance domains, including action planning (Cooper & Shallice, 2000), music performance (Palmer & Pfordresher, 2003), speech production (Bohland, Bullock, & Guenther, 2009), spelling (Glasspool & Houghton, 2005), typing (Rumelhart & Norman, 1982), and, of course, verbal STM (Burgess & Hitch, 1999; Henson, 1998a; Page & Norris, 1998), suggesting that competitive queuing may be a general basis for all serial behaviors (Bullock, 2004; Bullock & Rhodes, 2003; Glasspool, 2005).

Within the verbal competitive queuing system—namely, the phonological loop—Hurlstone et al. (2014) noted that there is direct evidence implicating a confluence of representational principles, including position marking, a primacy gradient, response suppression,

and output interference. However, they noted that the principles underlying the representation of serial order within the spatial competitive queuing system—namely, the visuospatial sketchpad—are less transparent, largely because the behavioral signatures of the aforementioned principles have not yet been studied in the spatial domain. Indeed, as noted earlier, most studies of spatial STM have focused on serial position curves and transposition gradients, and—as we show by simulation shortly—these data patterns are explicable by various different mechanisms for representing serial order within the class of competitive queuing models.

In the current study, we test the possible involvement of the different representational principles in spatial STM by focusing on a data pattern that previous research has shown is particularly effective in differentiating their theoretical predictions. Farrell and Lewandowsky (Farrell & Lewandowsky, 2004; Lewandowsky & Farrell, 2008) employed a generic competitive queuing architecture to contrast the response probability and recall latency predictions of five models built using different combinations of the previously mentioned principles for representing serial order. They found that although the models could not be distinguished on the basis of their predicted serial position curves and transposition gradients, they nevertheless made dramatically different predictions concerning the response times accompanying transpositions.

In what follows, we sought to identify the principles underlying the representation of serial order in spatial STM by examining the pattern of transposition latencies underlying a spatial serial recall task. We proceed as follows: We begin by presenting the characteristic response probability and recall latency predictions of five alternative mechanisms for the representation of serial order. Next, we report the results of three experiments that tested their error latency predictions. Finally, we present quantitative fits of the models to representative data from our experiments before exploring the robustness of their behavior to variations in model parameter settings.

## Modeling Transposition Latencies

Following Farrell and Lewandowsky (Farrell & Lewandowsky, 2004; Lewandowsky & Farrell, 2008), we did not utilize a fully implemented competitive queuing architecture for our simulations, but instead employed a single-layer lateral inhibition network, corresponding to the competitive choice layer. For each of the representational principles being modeled, we specified—using appropriate parameters—the profile of activations that would be expected initially at each output position in the parallel planning layer, before feeding that pattern of activations into the lateral inhibition network in order to generate an unambiguous response and an associated recall latency. Thus, we did not simulate the process of encoding serial order, as the selection mechanism is insensitive to the exact mechanisms generating the initial activations used to drive recall.

## A Common Competitive Queuing Response Selection Network Architecture

A schematic of the response selection network employed for the simulations is illustrated in Figure 2. It consists of a single competitive layer of localist item nodes corresponding to the pool of response elements from which sequences can be generated. Each
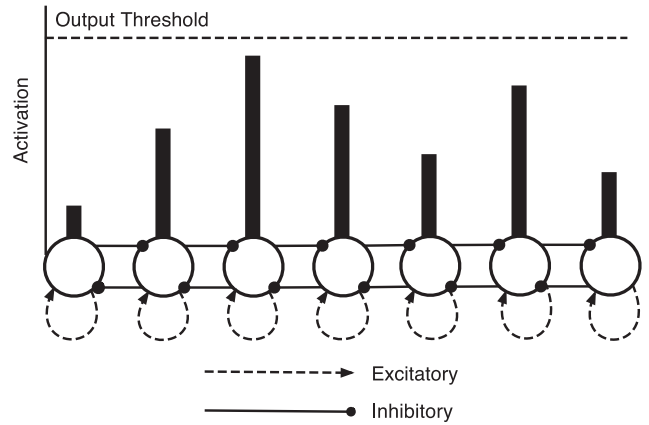


*Figure 2.* A schematic of the lateral inhibition neural network architecture employed for the simulations. The network is modeled on the competitive choice layer employed in competitive queuing models of serial behavior (Bullock, 2004; Bullock & Rhodes, 2003). Each localist item node possesses a single recurrent excitatory connection, as well as lateral inhibitory connections to all other item nodes. Nodes are fully interconnected, but only adjacent-neighbor inhibitory connections are shown to prevent visual cluttering. The activation of a node within the layer is determined by the external input it receives from outside the layer, plus positive recurrent self-excitation, and within-layer lateral inhibition. The number of nodes in the network is dependent on the sequence length being modeled.

node has a recurrent self-excitatory connection, plus lateral inhibitory connections to all other nodes. The excitatory and inhibitory weights are a hardwired property of the network and were set to constant values of 1.1 and $-0.1$, respectively. This network operates as a competitive filter that selects a single response from among a set of parallel activated representations. As noted earlier, serial order is represented in the network by setting starting activation values for the item nodes at each output position, the derivation of which is determined by the representational principles being modeled (see later). The activation of each node is determined by this initial external input, plus recurrent self-excitation, and lateral inhibition received from all other item nodes, which are jointly determined by the following equation (Houghton, 2005):

$$Int_{j,t} = a_{j,t-1} W^+ + W^- \sum_{i \neq j} a_{i,t-1} \quad (1)$$

where $Int_j$ constitutes the internal input a unit receives from within the layer, $a_j$ represents the initial activation of unit $j$ determined by its external input, $a_i$ constitutes the activation of all other nodes in the layer, $W^+$ and $W^-$ represent the excitatory and inhibitory weight values, respectively, and $t$ corresponds to time (note that negative activations are not allowed to spread in Equation 1; otherwise a node with a negative activation would send excitation, rather than inhibition, to other nodes within the layer). The first term on the right-hand side of Equation 1 represents the recurrent self-excitation, whereas the second term represents the lateral inhibition received from all other nodes in the layer. This sets up a winner-takes-all response competition over the item nodes, and the initially most active node—the node receiving the highest external activation—has the advantage that it will send more activation to itself than any other node and will also receive the least lateral inhibition. The node activations are iteratively updated

over time using Equation 1. This results in a gradual increase in the activation of the strongest node, and a gradual decrease in the activations of the weaker nodes as they receive more lateral inhibition. The iterations stop when the strongest node exceeds a response threshold $T$ (set to 1.0 for all simulations), and the number of iterations required to determine the response is taken as the network's recall latency. This process is repeated at each successive output position by defining a new set of starting activations and allowing activations to iterate toward a response. In order to bring the predicted recall times of the network within the range of the observed latencies in the forthcoming experiments, they were multiplied by a scaling parameter $S$ ($0 < S \leq 200$; where $S = 50$ for the initial simulations—1 iterative cycle = 50 ms).

Order errors are introduced by adding a small amount of random Gaussian noise ($\mu = 0$ and $\sigma = .04$ for the initial simulations) to the node activations on each iterative cycle. There is no temporal deadline for the network to converge on a response, which means that omission errors are not possible, nor are extralist intrusion errors, because the only nodes to receive activation are those corresponding to items that are part of the sequence.
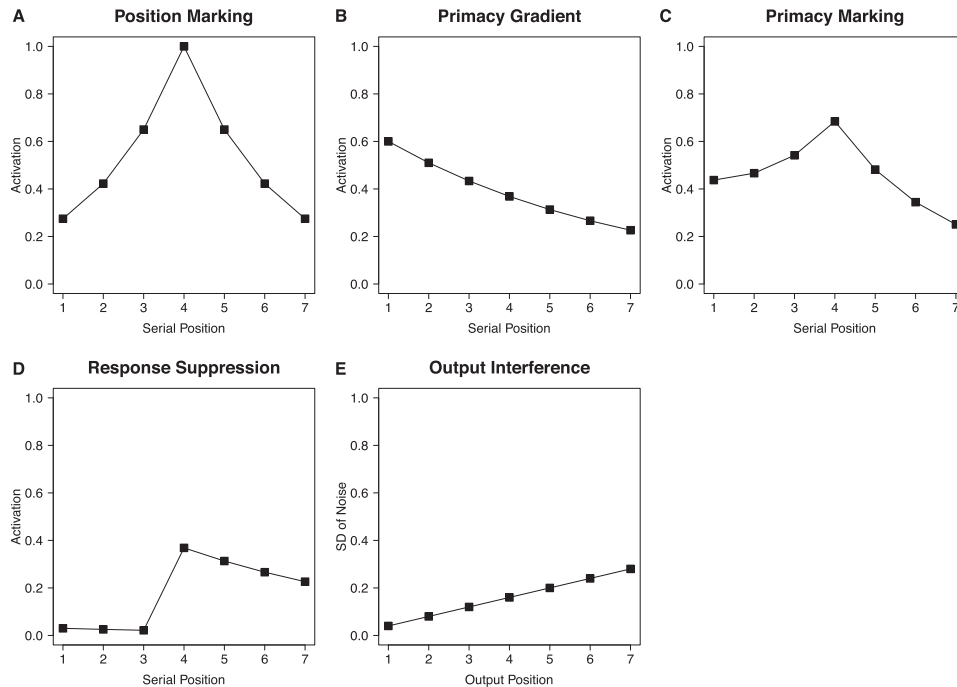
## Implementation of Representational Principles

The representational principles were implemented through different settings of the starting activations at each output position. The representational principles were instantiated as follows.

**Position marking.** Position marking was implemented by specifying activations for item nodes that reflected the distances between item positions. Specifically, the activation $a$ of the item node $j$ for the current response (output) $r$ position was strongest, whereas the activations of neighboring item nodes decreased as an accelerating function of their distance from the target item:

$$a_j = \lambda\theta^{|j-r|} \tag{2}$$

where $\theta$ is a parameter controlling the distinctiveness of the position marking activations ($0 < \theta \leq 1$; $\theta = .65$ for the initial simulations) and $\lambda$ is a weighting parameter that determines the distance of each item's initial activation from the response threshold ($0 < \lambda \leq 1$; $\lambda = 1$ for the initial simulations). For each output position, the activations generated by $\theta$ were rescaled to sum to 1—calculated by dividing each node's activation by the sum of the activations of all nodes—before they were multiplied by $\lambda$. This representational scheme produces gradients of activations akin to those generated by the positional context signals in the Burgess and Hitch (1999) and OSCAR (Brown et al., 2000) models. Figure 3A shows example starting activations for position marking for the fourth output position in a seven-item sequence.



*Figure 3.* Example starting activations for the four representational principles and combination of principles for seven-item sequences, based on the parameter settings employed for the initial simulations: (A) position marking (showing activations for the fourth output position); (B) a primacy gradient (showing activations for the first output position); (C) a primacy gradient and position marking (showing activations for the fourth output position); (D) response suppression (showing activations for a primacy gradient with suppression of the first three recalled items); (E) output interference (showing the increase in Gaussian noise applied to the starting activations across output positions). Panels A and C show activations for position marking without the rescaling described in the main text.

**Primacy gradient.** The primacy gradient was implemented as a decrease in activations across input positions. The activation of each node was determined by

$$a_j = a_1 \gamma^{j-1} \qquad (3)$$

where $a_1$ is the activation of the item node corresponding to the first input position ($0 < a_1 \leq 1$; $a_1 = .6$ for the initial simulations), and $\gamma$ is a parameter controlling the steepness of the primacy gradient ($0 < \gamma \leq 1$; $\gamma = .85$ for the initial simulations). Retrieval commenced by imposing the entire primacy gradient over the item nodes at the first output position and allowing activation to accumulate toward a response. This process was then repeated for each subsequent output position by imposing the same primacy gradient over the item nodes but with suppression (see later) of those nodes corresponding to previously recalled items. Example starting activations for a primacy gradient for the first output position are shown in Figure 3B.

**Primacy gradient and position marking.** In line with the seriating mechanisms instantiated in several theories of serial recall (Brown et al., 2000; Burgess & Hitch, 1999; Lewandowsky & Farrell, 2008), in some simulations, serial order was represented through the combination of a primacy gradient and position marking by calculating starting activations as follows:

$$a_j = (1 - \omega) a_1 \gamma^{j-1} + \omega \lambda \theta^{|j-r|} \qquad (4)$$

Equation 4 integrates Equations 2 and 3, and incorporates an additional weighting parameter $\omega$ ($0 < \omega \leq 1$; $\omega = .5$ for the initial simulations) that governs the relative importance of the two representations of serial order. When $\omega = .5$, the two representations of order are weighted equally. However, when $\omega < .5$, more weight is given to the primacy gradient representation of order; conversely, when $\omega > .5$, more weight is given to the positional representation of order. Figure 3C shows example starting activations for the combination of a primacy gradient and position marking for the fourth output position.

**Response suppression.** Response suppression was implemented by reducing an item's activation once it had been recalled. For each output position, starting activation values were first calculated based on the other representational principles being modeled. The activations of nodes corresponding to items that had already been recalled were then multiplied by $1 - \alpha$, where $\alpha$ represents the extent of response suppression ($0 < \alpha \leq 1$; $\alpha = .95$ for the initial simulations). Example starting activations for a primacy gradient complemented by response suppression for the fourth output position are illustrated in Figure 3D.

**Output interference.** Output interference was modeled by assuming that recall of an item added noise to the activations of yet-to-be-recalled items. Accordingly, random Gaussian noise with a standard deviation that increased as a function of output position was applied to the starting activations generated by the serial ordering principle(s) being modeled (e.g., position marking), and was determined by $\delta \times \sigma \times r$, where $\delta$ is a parameter controlling the weighting of output interference across output positions ($0 < \delta \leq 1$; $\delta = .5$ for the initial simulations) and $\sigma$ is the standard deviation of noise applied to activations during the iterative updating process (see earlier). An example of the increase in the standard deviation of Gaussian noise applied to the starting activations across output positions is shown in Figure 3E.

## Five Models of Serial Order

The response probability and recall latency predictions of five models of serial order—built from different combinations of the four principles—were compared: (a) position marking (PM), (b) position marking and response suppression (PM + RS), (c) position marking and output interference (PM + OI), (d) a primacy gradient and response suppression (PG + RS), and (e) a primacy gradient, position marking, and response suppression (PG + PM + RS). These models are the same as those employed by Farrell and Lewandowsky (Farrell & Lewandowsky, 2004; Lewandowsky & Farrell, 2008), and are representative of the range of mechanisms instantiated in contemporary theories of serial recall (see Table 1). Predictions were generated for each model using 50,000 simulation trials of seven-item sequences.

Figure 4 shows the theoretical predictions of the models for four serial recall measures: (a) accuracy serial position curves, (b) transposition gradients, (c) latency serial position curves, and (d) latency-displacement functions (LDFs). These predictions can be compared with those presented in Figures 3 and 4 of Farrell and Lewandowsky (2004), and Figure 1 of Lewandowsky and Farrell (2008). Considering, first, the accuracy serial position curves (Figure 4A), it can be seen that all models predict both a primacy and a recency effect. With the exception of the PM model—which predicts a symmetrical curve—the models correctly predict a more extensive primacy than recency effect. From inspection of the associated model predictions for transpositions (Figure 4B), it can be seen that all five models reproduce the three hallmark features of transposition gradients delineated at the outset, namely, the peaking of the gradients at displacement zero, the locality constraint, and the approximate symmetry of the anticipation and postponement error gradients. However, the PM + RS and PM + OI models predict somewhat more asymmetric transposition gradients, with postponements being slightly more frequent than anticipations. Turning to the predicted latency serial position curves (Figure 4C), all models correctly predict that recall latency varies inversely with recall accuracy. However, the models miss out on the extra long initial recall latency observed empirically. It has been suggested that this long initial output time reflects the operation of a preparatory stage that precedes production of the first response (Anderson & Matessa, 1997), such as the priming of a low-level motor output buffer (Sternberg, Monsell, Knoll, & Wright, 1978). Because this preparatory latency does not assist in discriminating between the models—and because ancillary assumptions are necessary to accommodate it—no attempt is made to model it here.

It is apparent from the foregoing that the five models generate qualitatively similar predictions for the first three serial recall measures, and that these measures therefore cannot be used to discriminate between them. However, adjudication becomes possible when one considers their predicted LDFs, illustrated in Figure 4D. To elaborate, LDF plots have a similar form as transposition displacement gradients, but with mean latency, rather than response probability, as the dependent measure. Following Farrell and Lewandowsky (2004), the effect of output position has been removed from the model LDFs by subtracting their predicted mean recall latencies at each output position from the individual latencies at corresponding positions. This filtering process is necessary because output position is correlated with transposition displace-
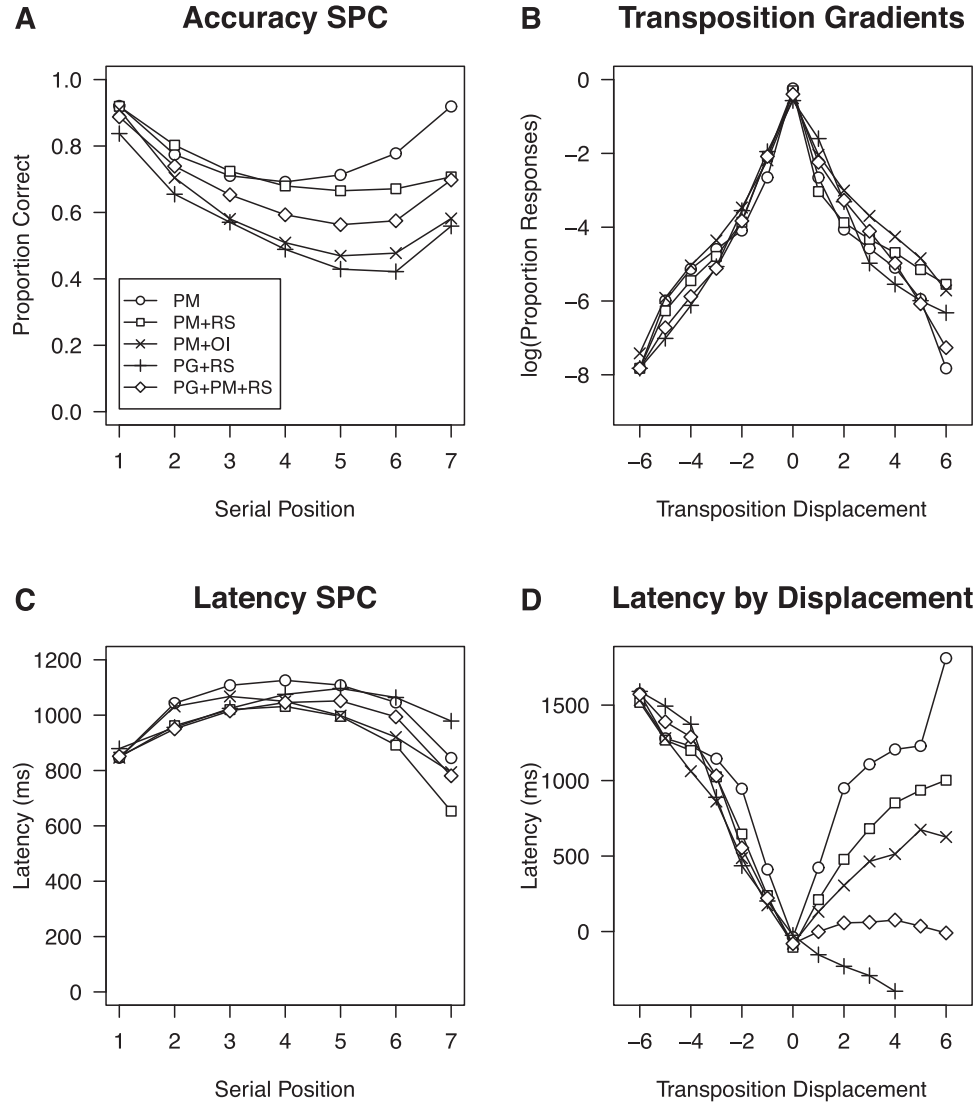
*Figure 4.* Response probability and recall latency predictions for five models of serial order. Panels show accuracy serial position curves (A), transposition gradients (B), latency serial position curves (C), and latency-displacement functions (D). PM = position marking; RS = response suppression; OI = output interference; PG = primacy gradient.

ment: Most anticipations occur near the beginning of the sequence, whereas most postponements occur toward the end. This is problematic because recall latencies are typically longer at early, than at late, output positions, which can artificially inflate the recall latencies for anticipations and artificially accelerate the recall latencies for postponements. The filtering process removes any effects of output position from the data, permitting an uncontaminated examination of the effects of transposition displacement on the dynamics of transpositions (note that the negative latencies at some transposition displacements are a consequence of this filtering). Before scrutinizing the model LDFs, one further issue warrants comment. Whereas repetition errors—both occurrences of the repeat—were included in the predicted transposition gradients shown in Figure 4B, they were excluded from the associated LDFs in Figure 4D, because the models incorporating response suppres-

sion predict that the error latencies for repeated and nonrepeated items behave differently. Specifically, because suppression renders an already-recalled item a weak recall competitor at subsequent output positions, it will take some time for a suppressed item's activation to overcome the strong lateral inhibitory signals received from other items to win the output competition a second time. Thus, the models incorporating response suppression necessarily predict longer recall latencies for transpositions involving repeated items than transpositions involving nonrepeated items. To maintain consistency with the modeling, in the experiments that follow, repetitions were included in the transposition gradients but excluded from the accompanying LDFs.

The first thing to note about the LDFs is that all models predict that the slope of the function for anticipations is negative, indicating that latencies for anticipations increase as a function of dis-

placement. This is because whether serial order is represented by a primacy gradient, position marking, or by both principles, an anticipation involves a weakly activated item being recalled from among a set of stronger competitors from earlier input positions. The further an item is anticipated, the greater the number of competitors it must overcome and, consequently, the longer it will take for that item to win the output competition. It follows that the models cannot be distinguished on the basis of their predicted anticipation slopes. However, it is clear from inspection of Figure 4D that there is considerable heterogeneity between the models in their predicted postponement slopes. Starting with the PM model, this model predicts that latencies for postponements increase as a function of displacement in the same way as they do for anticipations, giving the overall LDF a symmetric V-shaped function. The PM + RS and PM + OI models both predict a partially asymmetric V-shaped LDF in which the slope of the function for postponements is reduced compared to that for anticipations. In the PM + RS model, this arises because the suppression of items following recall reduces the number of competitors at late output positions. Because postponements will tend to occur toward the end of the sequence, the reduced competitor set results in shorter recall latencies for these errors. In the PM + OI model, it arises because the impact of accumulating output interference is to gradually raise the activations of later items, taking them closer to the decision threshold. In sharp contrast to the preceding models, the PG + RS model predicts a negative postponement slope in which the latencies for postponements become accelerated with increasing positive displacements. This is because, in this model, a postponement error involves a strongly activated item being recalled from among a set of weaker competitors from later input positions. The longer an item is postponed, the greater the disparity will be between its activation and that of its weaker rivals, enabling it to quickly suppress those items through lateral inhibition and win the output competition. Finally, the PG + PM + RS model predicts a flat postponement slope because the impact of adding the primacy gradient to the position marking activations is to compress the component of the position marking activations representing positional uncertainty with respect to the beginning of the sequence.

These predictions are consistent with those reported by Farrell and Lewandowsky (Farrell & Lewandowsky, 2004; Lewandowsky & Farrell, 2008). These authors reported three experiments involving the keyboard-timed serial recall of sequences of letters and digits (Farrell & Lewandowsky, 2004), in which they consistently observed that transposition latency is a negative function of transposition displacement, with a reduction in the slope of the LDF for postponements, compared with anticipations—a pattern uniquely consistent with the theoretical prediction of the PG + PM + RS model. However, this model was not included in the original model comparisons of Farrell and Lewandowsky (2004), who—based on quantitative fits of the other four models to data from their Experiment 3—initially interpreted their results as conferring support for the PG + RS model. The PG + PM + RS model was only introduced in a subsequent review article by Lewandowsky and Farrell (2008), within which the authors presented qualitative predictions of the model, but no attempt was made to fit it to the same data employed for the original model comparisons. Hurlstone (2010) has plugged this theoretical gap by showing that when fit to

those same data, the PG + PM + RS model does indeed provide a better description of the observed LDF than the PG + RS model.

## Outline of Experiments

We now report three experiments that tested the transposition latency predictions of the five models employing a computerized version of the Corsi-Blocks Task. Unlike conventional versions of this task, the first two experiments employed a sequential—as opposed to a simultaneous—presentation format. To explain, in typical instantiations of the Corsi-Blocks Task (e.g., De Lillo, 2004; Fischer, 2001; Smyth & Scholey, 1996), the locations are always simultaneously visible and the presentation order of the sequence is indicated by highlighting each location in turn. The sequential presentation format adopted here involved displaying locations in isolation by having each briefly appear and then disappear in succession. This presentation format was chosen to increase correspondence with the verbal serial recall task, which uses a sequential presentation format. To ensure that the pattern of the LDFs observed in these initial experiments was not specific to the presentation method employed, a third experiment used the conventional simultaneous presentation format. In addition, Experiments 1 and 3 incorporated a temporal grouping manipulation, whereas Experiment 2 incorporated a distractor manipulation.

## Experiment 1

The first experiment examined the LDFs underpinning ungrouped and temporally grouped spatial sequences. Differentiating a sequence into subgroups by inserting temporal pauses has been shown in verbal studies to exert a multiplicity of effects on serial recall. First, grouping enhances overall recall accuracy (Frankish, 1985, 1989; Hitch, Burgess, Towse, & Culpin, 1996; Ryan, 1969a, 1969b) and modifies the shape of the accuracy serial position curve: Primacy and recency effects are observed within groups, as well as the sequence overall (Frankish, 1985, 1989; Hitch et al., 1996). Second, grouping exerts systematic effects on the pattern of recall latencies: In addition to leaving a long pause prior to outputting the first item in the sequence, participants leave long pauses prior to outputting the first item of each group (Farrell & Lelièvre, 2009; Farrell & Lewandowsky, 2004; Maybery et al., 2002; Parmentier & Maybery, 2008). Third, although grouping reduces the incidence of transpositions overall, it increases the incidence of *interpositions*—transpositions between groups that maintain their positions within groups (Farrell & Lelièvre, 2009; Farrell & Lewandowsky, 2004; Henson, 1996, 1999; Ng & Maybery, 2002, 2005; Ryan, 1969a, 1969b).

There is a general consensus among serial-recall theorists (see, e.g., Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998a; Lewandowsky & Farrell, 2008) that temporal grouping effects confer evidence for the operation of position marking in verbal STM. Positional models account for such effects by assuming that order information in grouped sequences is represented on multiple dimensions, with one dimension representing the positions of groups (Brown et al., 2000; Henson, 1998a; Lewandowsky & Farrell, 2008) or items (Burgess & Hitch, 1999) in the sequence overall, and with a second dimension representing the positions of items within groups. The latter dimension is crucially necessary to account for the interposition errors that are a hallmark characteristic of grouped sequences.

Temporal grouping effects are not confined to verbal memoranda; they have also been documented in a study by Parmentier and colleagues (2006; Experiment 4), in which participants reproduced the order of sequences of seen spatial locations. Interestingly, however, unlike in the verbal studies of temporal grouping, Parmentier and colleagues failed to observe an increase in the incidence of interpositions in grouped spatial sequences. This discrepancy is noteworthy because the occurrence of interpositions in grouped sequences is a core component of the claim that the representation of serial order in such sequences is based on a process of position marking (Henson, 1999).

Given its multifarious effects on recall, we incorporated the temporal grouping manipulation to establish whether it exerted any systematic effects on the dynamics of transpositions. Specifically, if grouping induces a strong reliance on positional representations, the postponement slope of the LDF for grouped sequences should be steeply positive—yielding an overall symmetric or partially asymmetric V-shaped LDF—consistent with the predictions of positional models. However, if the postponement slope is flat or negative, then this would point to a role for principles other than position marking in the representation of serial order in grouped spatial sequences (viz., a primacy gradient with response suppression). An ancillary reason for incorporating the grouping manipulation was to establish the generality of the results of Parmentier et al. (2006)—notably, their failure to witness an increase in interpositions in grouped spatial sequences.

Although studies of temporal grouping in the verbal domain have typically employed six-item sequences organized into two groups of three (Farrell, 2008; Farrell & Lewandowsky, 2004; Maybery et al., 2002; Parmentier & Maybery, 2008), this method was not adopted here because piloting revealed that recall accuracy was too high. Instead, seven-item sequences were employed, and temporally grouped sequences were divided into a group of four items followed by a group of three—a grouping pattern employed previously in some verbal studies (Farrell & Lelièvre, 2009; Henson, 1999). One implication of employing groups of unequal sizes is that interpositions can fall into two types: absolute and relative. Absolute interpositions are transpositions between groups that maintain their absolute position within groups, whereas relative interpositions are transpositions between groups that maintain their relative position within groups. Using a 4–3 grouping pattern, absolute interpositions are reflected by ±4 displacements, whereas relative interpositions are reflected by ±3 displacements involving terminal group items. We take an increase in the incidence of either of these errors as evidence of within-group positional coding in spatial STM.

## Method

**Participants.**    Twenty members of the campus community at the University of York took part in the experiment in exchange for course credit or payment of £4 (approximately $6.50).

**Stimuli and apparatus.**    The stimuli were sequences composed of seven spatial locations. The locations were nine fixed gray icons (measuring 5 cm × 5 cm each) arranged haphazardly on a white background (see Figure 5). The minimum and maximum distances between pairs of locations (measured from the center of each square) were 9.2 cm and 35.3 cm, respectively. Stimulus presentation and response collection were controlled via software
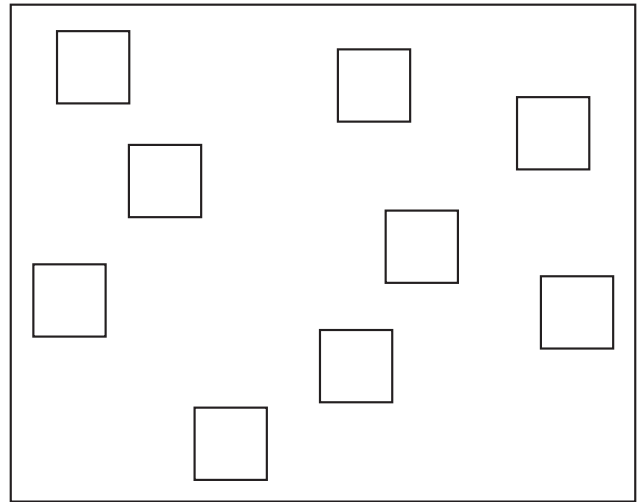


*Figure 5.*    Schematic of the placement of the nine squares from which items were randomly selected for sequential presentation in the spatial serial recall task.

developed in-house using a Dell Optiplex (Intel Core 2 Duo, 2.13 GHz processor) PC equipped with a 19-in. monitor and a Razer Copperhead high-precision mouse. The same apparatus was used for all subsequent experiments.

**Design and procedure.**    The experiment manipulated a single independent variable: sequence type (ungrouped vs. grouped), which was a within-subjects factor. Participants always received the grouped sequences last in order to reduce the likelihood that they would subjectively group the ostensibly ungrouped sequences.

Participants were tested individually in a quiet room in the presence of the experimenter. They initiated each trial by selecting a "begin trial" icon located in the center of the computer display using the mouse-driven pointer. Following a 1,000-ms blank interval, any seven of the nine possible locations were displayed individually on-screen in a random order. For ungrouped sequences, locations were presented for 500 ms each, with a 250-ms blank interstimulus interval. For grouped sequences, the presentation rates were the same, except that the interstimulus interval separating presentation of the fourth and fifth locations was 1,000 ms, creating the impression of two groups of locations, the first containing four locations and the second containing three.

Following the final location, there was a 1,000-ms blank interval, after which all nine locations appeared simultaneously on-screen, prompting participants to recall the sequence in forward serial order using the mouse-driven pointer. Once an item was selected, its color changed transitorily from gray to green for 50 ms to acknowledge that the response had been registered by the computer. Locations could be selected on more than one occasion, meaning that repetition errors were possible, as were intrusion errors, because the spatial array included the two locations that had not been presented in the to be remembered sequence. Following each response, a counter located in the bottom right-hand corner of the screen incremented in value to inform the participant of the number of responses executed so far. Participants were encouraged to guess if they were uncertain of the location for a given position, but if no location came to mind, a "don't know" response could be

registered by selecting a question mark, which was located adjacent to the response counter. Once seven responses had been recorded by the computer, the display cleared and the recall time for the sequence was conveyed in the central screen position for 3,000 ms, after which it was replaced by the "begin trial" icon for the subsequent trial.

Participants were instructed to encode sequences visually, without deploying supplementary verbal or gestural encoding strategies. All participants reported compliance with these instructions. The experiment consisted of two practice and 80 experimental trials for each sequence type. Sessions lasted approximately 60 min.

## Results

The data were analyzed using a strict serial recall scoring procedure: An item was only scored as correct if its output serial position was the same as its input serial position. The results are structured into four sections: (a) accuracy serial position curves, (b) transposition gradients, (c) latency serial position curves, and (d) LDFs. Effect size estimates are provided—for focused comparisons only—using Pearsons $r$.

**Accuracy serial position curves.** The accuracy serial position curves are illustrated in Figure 6A, from which it can be seen that grouping enhanced overall recall accuracy and caused a change in the shape of the serial curve: Primacy and recency effects can be seen within groups, as well as the sequence overall. Statistical confirmation of these observations was obtained by means of a 2 (sequence type) $\times$ 7 (serial position) ANOVA,[2] which revealed significant main effects of sequence type, $F(1, 19) = 8.45$, $p < .01$, $r = .55$, and serial position, $F(6, 114) = 59.81$, $p < .001$, together with a significant interaction, $F(6, 114) = 11.26$, $p < .001$.

**Transposition gradients.** The transposition gradients are shown in Figure 6B and exhibit the three hallmark characteristics delineated earlier. If grouping had promoted an increase in absolute or relative interpositions, then peaks in the transposition gradients for grouped sequences at displacements $\pm 4$ and $\pm 3$, respectively, should be apparent. However, with the exception of a slight elevation in responses at displacement $-4$, there is little indication that grouping engendered an increase in such interpositions. To scrutinize the error patterns more closely, transpositions were classified as occurring within or between groups, with the latter errors being further subdivided into interpositions and position nonpreserving errors. Because the incidence of interpositions was low, absolute and relative interpositions were combined. Paired comparisons performed on the log-odds transformed error proportions revealed that grouping increased the incidence of transpositions within groups (ungrouped $M = .71$ vs. grouped $M = .87$), $t(19) = -6.94$, $p < .001$, $r = .85$, but decreased the incidence of position nonpreserving transpositions between groups (ungrouped $M = .24$ vs. grouped $M = .09$), $t(19) = 7.4$, $p < .001$, $r = .86$. Critically, grouping did not increase the incidence of interpositions (ungrouped $M = .05$ vs. grouped $M = .04$), $t(19) = 1.27$, $p = .220$, $r = .28$.

**Latency serial position curves.** The mean recall latencies associated with correct responses at each serial position are portrayed in Figure 6C. It can be seen by inspection that the latency curve for ungrouped sequences exhibits a long output time for the first position, with the remainder of the curve following an inverted-U-shaped profile. By contrast, the latency curve for grouped sequences is relatively flat but with discrete peaks at Serial Positions 1 and 5 corresponding to the recall of the first item of each group. Reflecting these trends, a 2 (sequence type) $\times$ 7 (serial position) ANOVA performed on the data revealed significant main effects of sequence type, $F(1, 19) = 33.49$, $p < .001$, $r = .80$, and serial position, $F(6, 114) = 12.45$, $p < .001$, together with a significant interaction, $F(6, 114) = 3.781$, $p < .05$.

**Latency-displacement functions.** Turning to the data of chief interest, Figure 6D shows the LDFs, which plot the mean recall latencies of transpositions as a function of transposition displacement.[3] Note that the effect of output position has been removed from these data in the same way as in the model predictions by taking each participant's mean recall latency at each output position and subtracting it from the individual latencies at the same position. The reader is reminded that the negative latencies at some transposition displacements are a consequence of this filtering process.

It is apparent from inspection of Figure 5D that with the exception of some unique deviations—the accelerated recall latencies at displacement $-5$ in ungrouped sequences and $-6$ in grouped sequences—the slopes of the functions for anticipations are negative, with a slightly shallower slope for grouped than for ungrouped sequences. By contrast, the slopes of the functions for postponements are approximately flat, albeit with a slight acceleration in the recall latencies at displacements $+5$ and $+6$. The LDFs were analyzed using the following two-stage procedure. In the first stage, regression analyses were performed that examined the relationship between transposition latency and transposition displacement for each individual participant. One analysis examined the relationship between latency and displacement for anticipations (displacements $-6$ to $0$), and a second examined the relationship between latency and displacement for postponements (displacements $0$ to $+6$). Thus, regression equations were computed for each participant by regressing transposition latency on displacements that were anticipations and postponements, separately.

In the second stage, the regression parameter estimates for the slopes of the LDFs for anticipations and postponements were pooled together and subjected to one-sample $t$ tests to determine whether they deviated reliably from zero. The regression statistics for the slope analyses are summarized in Table 2, from which it can be seen that the mean parameter estimates for the slopes of the

---

[2] For all within-subjects ANOVAs reported in this article, violations of the assumption of sphericity were accommodated by use of the Greenhouse-Geisser correction.

[3] Exploratory analyses of the LDFs for all three experiments reported in this article were performed initially to examine their sensitivity to potential outliers. These analyses revealed that the qualitative form of the LDFs (the sign and steepness of the slopes of the functions for anticipations and postponements) was generally unaffected by whether all latencies were included in the analysis, or only those latencies within the range of 2.5 or 3 standard deviations from the mean. The empirical pattern was also similar when the mean of the median latencies was used as the dependent measure. Accordingly, given the insensitivity of the qualitative form of the LDFs to these different methods of dealing with response time outliers, we retained all of participants' responses, in order to maximize the number of errors available for the transposition latency analyses, and we used the mean latency as the dependent measure.
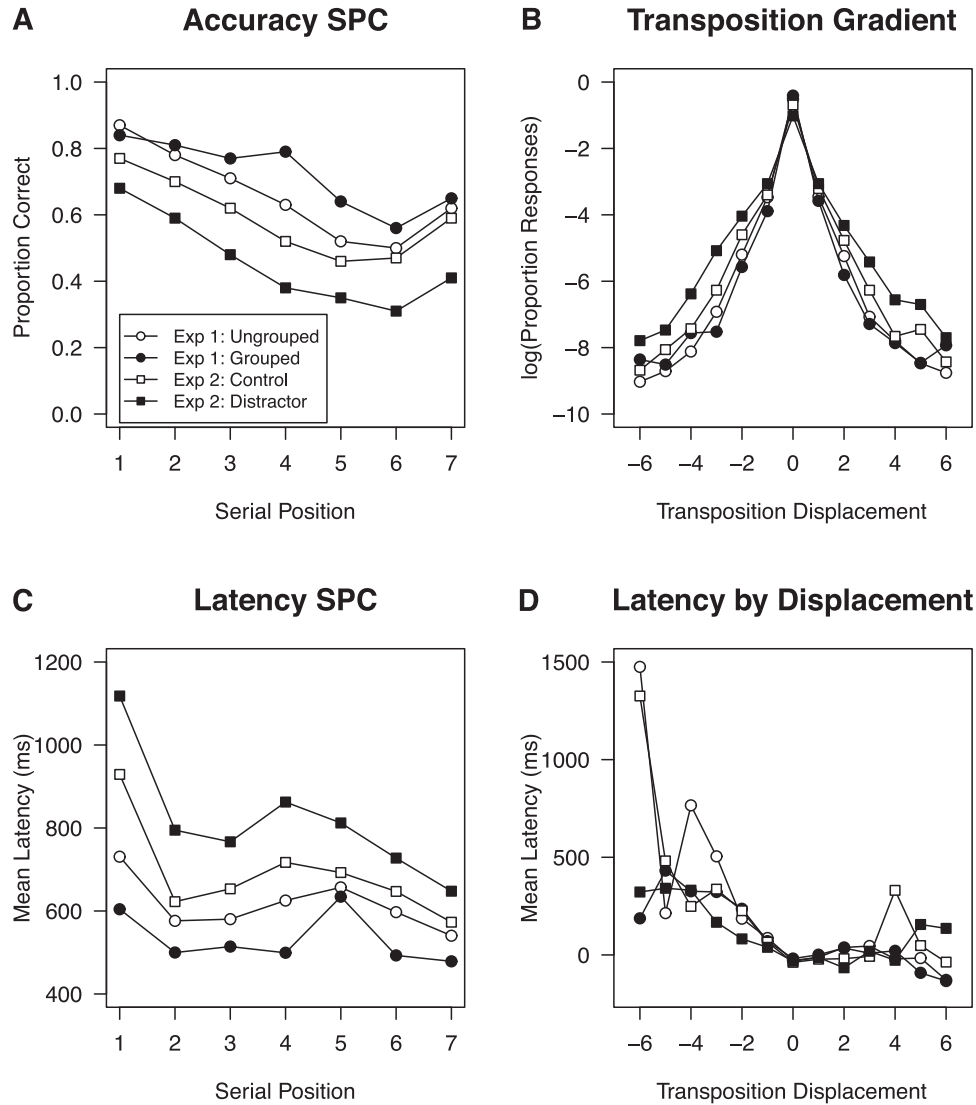
*Figure 6.* Response probability and recall latency data for Experiments 1 and 2. Panels show accuracy serial position curves (A), transposition gradients (B), latency serial position curves (C), and latency-displacement functions (D).

functions for anticipations were negative and deviated reliably from zero: $t(19) = -3.68$, $p < .01$, $r = .64$, for ungrouped sequences, and $t(19) = -3.09$, $p < .01$, $r = .58$, for grouped sequences. In contrast, the mean parameter estimates for the slopes of the functions for postponements did not differ significantly from zero: $t(19) = -.48$, $p = .636$, $r = .14$, for ungrouped sequences, and $t(19) = .25$, $p = .808$, $r = .06$, for grouped sequences.

To give some indication of the variability in LDF slopes, Figure 7 shows the anticipation and postponement slope estimates for individual participants. It is evident by inspection that for both sequence types, the majority of participants contributed steep negative slopes for anticipations, whereas for postponements, there was an approximately equal distribution of negative and positive postponement slopes, but these were exclusively shallow slopes. This confirms that the empirical pattern of the aggregate LDFs is an accurate reflection of the individual LDFs from which they are composed, and not the result of a small number of individuals exerting undue influence on the data.

## Discussion

Although the findings of central interest are the LDFs, we begin by considering the impact of the temporal grouping manipulation on the other recall measures. Consistent with the spatial memory study of Parmentier et al. (2006), as well as several verbal studies (Farrell & Lelièvre, 2009; Farrell & Lewandowsky, 2004; Henson, 1999; Hitch et al., 1996; Maybery et al., 2002), temporal grouping enhanced overall recall accuracy, caused a change in the shape of the accuracy and latency serial position curves, and reduced the probability of transpositions between groups. These findings buttress the hypothesis that positional information contributes to the encoding of spatial sequences, and that when such sequences are

Table 2

*Mean Regression Parameter Estimates for the Slopes of the Latency-Displacement Functions for Anticipations and Postponements in Experiments 1 and 2*

| Parameter | Estimate | SE | t | p | r |
|---|---|---|---|---|---|
| Experiment 1 | | | | | |
| Ungrouped | | | | | |
| Anticipation | −169.22 | 46.04 | −3.68 | .00 | .64 |
| Postponement | −4.07 | 8.48 | −.48 | .64 | .11 |
| Grouped | | | | | |
| Anticipation | −120.66 | 39.03 | −3.09 | .00 | .58 |
| Postponement | 1.80 | 7.30 | .246 | .81 | .06 |
| Experiment 2 | | | | | |
| Control | | | | | |
| Anticipation | −152.06 | 44.37 | −3.43 | .00 | .62 |
| Postponement | 57.02 | 40.80 | 1.4 | .18 | .31 |
| Distractor | | | | | |
| Anticipation | −89.51 | 30.29 | −2.96 | .00 | .56 |
| Postponement | 30.47 | 16.92 | 1.8 | .09 | .38 |

temporally grouped, this positional information assumes a multidimensional form. Based on the hallmark finding in verbal studies that temporal grouping promotes an increase in the incidence of interpositions, it has been hypothesized that one of these dimensions represents the positions of items within groups (Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998a; Lewandowsky & Farrell, 2008). However, consistent with the results of Parmentier and colleagues (2006), the present experiment failed to observe an increase in the probability of such errors in temporally grouped spatial sequences. We consider the theoretical implications of this result in the General Discussion section, but for now, we note that a further attempt to replicate the pattern of interpositions documented in verbal studies of temporal grouping is reported in Experiment 3.

Turning to the findings of chief interest, the LDFs for ungrouped and grouped spatial sequences exhibited an overall negative trend, but with a flattening of the slopes of the functions for postponements compared with anticipations. This empirical pattern is consistent with that observed in the verbal serial recall experiments of Farrell and Lewandowsky (2004), and is most compatible with the error latency prediction of a representational mechanism combining a primacy gradient, position marking, and response suppression. Thus, although the effects of temporal grouping just described point to a pivotal role for position marking in the representation of serial order in spatial sequences, the LDFs point to a necessary role for a primacy gradient and response suppression, because combinations of position marking that do not invoke a primacy gradient predict steep positive postponement slopes that were not observed empirically. Indeed, the insensitivity of the postponement slope of the LDF to the temporal grouping manipulation—which was also observed in the experiments of Farrell and Lewandowsky (2004)—suggests that even temporally grouped spatial sequences recruit a primacy gradient.

## Experiment 2

One limitation of Experiment 1 was that insufficient errors were observed at some transposition displacements: The percentage of missing data cells for the LDF analyses was 36% for ungrouped

sequences and 41% for grouped sequences. The majority of these missing data cells represented long-range transpositions, specifically, displacements in the range of −4 to −6 and +4 to +6. It follows that a more reliable assessment of the LDFs can be obtained by increasing the number of observations for the transposition latency analysis. The aim of Experiment 2 was to increase the frequency of transpositions by incorporating an end-of-sequence distractor task. The distractor task involved making parity judgments to two digits that followed the presentation of each spatial sequence—a task employed previously to promote an increase in the frequency of transpositions (Farrell & Lelièvre, 2009).

### Method

**Participants.** Twenty members of the campus community at the University of York took part in the experiment in exchange for course credit or payment of £5 (approximately $8).

**Design and procedure.** The experiment manipulated a single independent variable: sequence type (control vs. distractor), which was a within-participants factor. Half of the participants received the control sequences followed by the distractor sequences, whereas the remaining half received the sequences in the converse order.
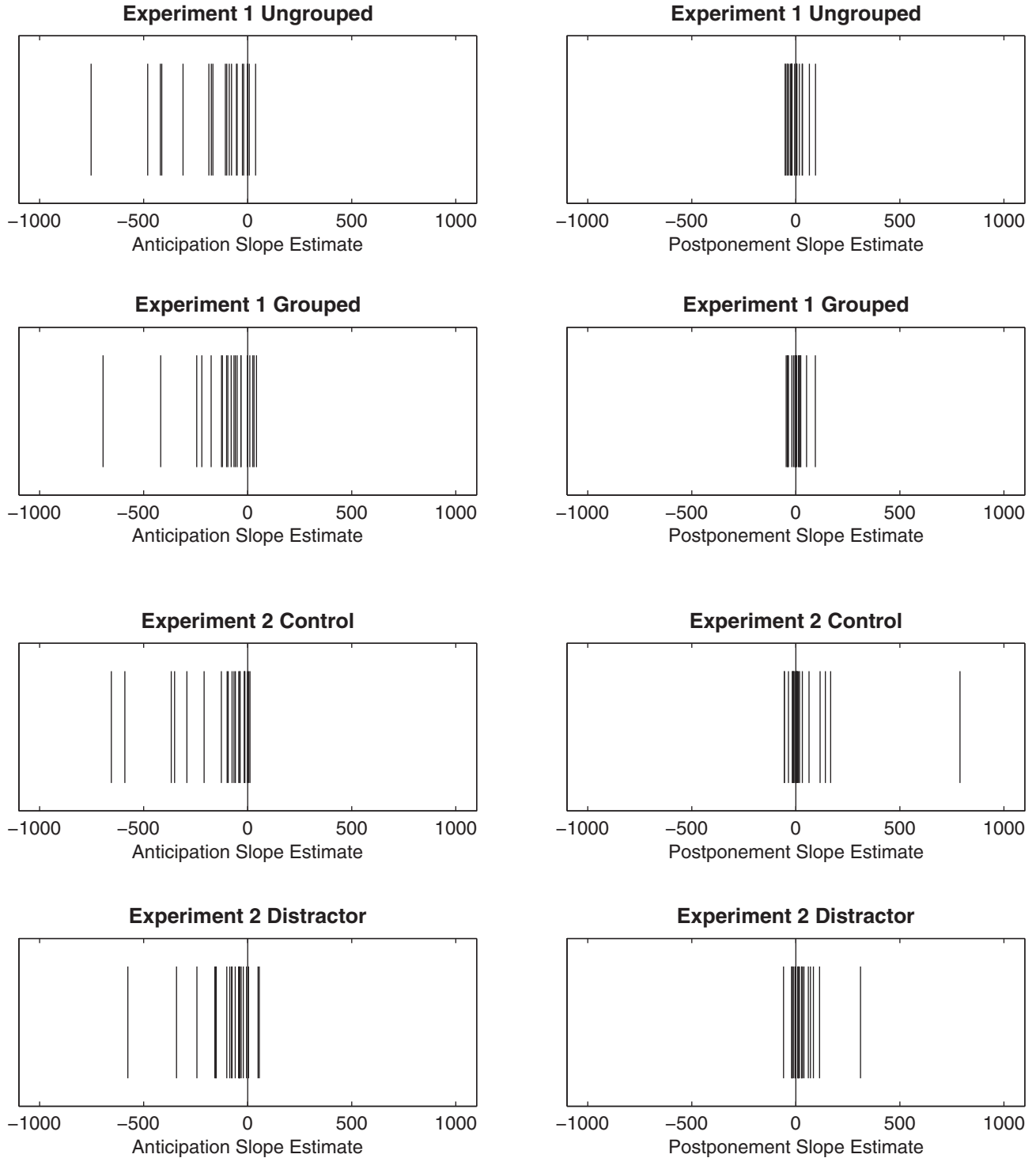
The procedure was identical to that for the ungrouped sequence condition of Experiment 1, with two exceptions. First, the interstimulus interval was increased from 250 ms to 500 ms. Second, for the distractor sequences, the 1,000-ms interval after presentation of the final location was followed by two digits presented individually in the central screen position. Participants were required to make parity judgments to each digit, pressing the left mouse button for odd digits and the right mouse button for even digits. Participants were informed that it was imperative that they classified each digit correctly. Following the parity judgment task, the locations appeared on-screen simultaneously, prompting participants to recall the sequence in forward serial order.

Participants were instructed to encode the sequences visually, without deploying supplementary verbal or gestural encoding strategies. All participants reported adherence to these instructions. The experiment consisted of two practice and 80 experimental trials for each sequence type. Sessions lasted approximately 75 min.

### Results

**Accuracy serial position curves.** Figure 6A shows the accuracy serial position curves, from which it can clearly be seen that the accuracy of recall was lower at all serial positions for distractor sequences than for control sequences—with the most marked effects being observed over the final two serial positions—confirming that the end-of-sequence distractors exerted the desired effect on performance (it is also apparent that recall accuracy for control sequences was appreciably lower than for corresponding ungrouped sequences in Experiment 1, for all but the last two serial positions). A 2 (sequence type) × 7 (serial position) ANOVA revealed significant main effects of sequence type, $F(1, 19) = 36.67$, $p < .001$, $r = .81$, and serial position, $F(6, 114) = 58.89$, $p < .001$, with the interaction falling marginally short of significance, $F(6, 114) = 2.54$, $p = .055$.

**Transposition gradients.** The transposition gradients are illustrated in Figure 6B and exhibit the usual features. Corroborating the serial position analysis, the incidence of anticipations and

**Experiment 1 Ungrouped**



**Experiment 1 Ungrouped**



**Experiment 1 Grouped**



**Experiment 1 Grouped**



**Experiment 2 Control**



**Experiment 2 Control**



**Experiment 2 Distractor**



**Experiment 2 Distractor**



*Figure 7.* Individual regression parameter estimates for the slopes of the latency-displacement functions (LDFs) for anticipations and postponements in Experiments 1 and 2. Each vertical line in a panel represents a regression slope estimate for a single participant, for a particular experimental condition. The left-hand panels show the estimates for the slopes of the LDFs for anticipations; the right-hand panels show the estimates for the slopes of the LDFs for postponements. The first and second rows of panels give the slope estimates for ungrouped and temporally grouped sequences, respectively, in Experiment 1; the third and fourth rows of panels give the slope estimates for control and distractor sequences, respectively, in Experiment 2. The full vertical line within each panel indicates a slope value of zero.

postponements was greater for distractor sequences than for control sequences.

**Latency serial position curves.** Figure 6C shows the latency serial position curves, which exhibit the usual pattern expected for ungrouped sequences: The recall times peak at the first output position, with the remainder of the curve following an inverted-U-shaped profile. Paralleling the accuracy serial position analysis, the distractor manipulation slowed down recall at all serial positions. Statistical confirmation of these trends was obtained by a 2 (sequence type) $\times$ 7 (serial position) ANOVA, which revealed significant main effects of sequence type, $F(1, 19) = 12.61$, $p < .01$, $r = .63$, and serial position, $F(6, 114) = 8.33$, $p < .01$, but no significant interaction, $F(6, 114) = 1.07$, $p = .387$.

**Latency-displacement functions.** The LDFs, with the effects of output position subtracted, are shown graphically in Figure 6D and exhibit the same empirical pattern documented in Experiment 1: The slopes of the functions for anticipations are negative, whereas with some unique deviations—the single peak at displacement $+4$ for control sequences and the peaks at displacements $+5$ and $+6$ for distractor sequences—the slopes of the functions for postponements are approximately flat. The regression statistics for the LDF slope analyses are shown in Table 2 and provide statistical confirmation of the pattern illustrated in Figure 6D. The slopes of the functions for anticipations were statistically negative: $t(19) = -3.43$, $p < .01$, $r = .62$, for control sequences, and $t(19) = -2.96$, $p < .01$, $r = .56$, for distractor sequences, whereas the slopes of the functions for postponements were weakly positive, but did not differ significantly from zero: $t(19) = 1.4$, $p = .178$, $r = .31$, for control sequences, and $t(19) = 1.8$, $p = .088$, $r = .38$, for distractor sequences. Inspecting the LDF slopes for individual participants (see Figure 7), it is apparent that for both sequence types, the majority of participants contributed steep negative slope estimates for anticipations, whereas for postponements, there was an approximately equal distribution of negative and positive postponement slopes, and with only two exceptions—one individual who contributed a steep positive slope in the control condition and another who contributed a steep positive slope in the distractor condition—these slopes were predominantly shallow slopes.

## Discussion

The requirement to make parity judgments to two digits presented at the end of spatial sequences significantly lowered the accuracy of recall and led to a corresponding increase in the frequency of transpositions, as was desired. Indeed, even the incidence of transpositions in control sequences was increased—relative to corresponding (ungrouped) sequences in Experiment 1—presumably because of the counterbalancing of conditions employed in the present experiment. The resulting LDFs were qualitatively consistent with those witnessed in the previous experiment, except that the slopes of the functions for postponements—although statistically flat—exhibited a slight positive trend. This was attributable to local peaks in the LDF at displacement $+4$ in control sequences and displacements $+5$ and $+6$ in distractor sequences, as opposed to a more general tendency for the recall times associated with postponements to slow down with increasing positive displacements. The empirically observed LDFs are once again most compatible with the error latency prediction of a representational

mechanism combining a primacy gradient, position marking, and response suppression.

## Experiment 3

The preceding experiments have consistently revealed that the LDFs associated with the recall of spatial sequences are characterized by negative anticipation slopes and flat postponement slopes. The generality of this empirical pattern is highlighted by its insensitivity to manipulations of temporal grouping and postsequence interference. The aim of Experiment 3 was to further examine the generality of this result in three ways. First, longer sequences of nine spatial locations were employed to further lower the accuracy of recall, thereby increasing the frequency of transpositions without incorporating end-of-sequence distractors. This also enabled an assessment of whether the relationship between transposition latency and transposition displacement hitherto observed holds when transpositions could span a larger range of displacements. Second, a simultaneous spatial presentation array was employed instead of a sequential presentation array, consistent with typical instantiations of the Corsi-Blocks Task. During presentation of the sequence, all nine locations were simultaneously visible and their presentation order was indicated by highlighting each location in turn. Third, a temporal grouping manipulation was incorporated once more, but this time employing the presentation format of grouping into threes, which has been the most widely employed grouping pattern in verbal studies (Farrell, 2008; Farrell & Lewandowsky, 2004; Frankish, 1989; Hitch et al., 1996; Maybery et al., 2002; Parmentier & Maybery, 2008; Ryan, 1969a, 1969b). An additional feature of Experiment 3 was that ungrouped and grouped sequences were administered to different groups of participants. This design choice was made to reduce the length of experimental sessions and to circumvent potential order artifacts that may have arisen in Experiment 1 because of the constant administration of ungrouped sequences prior to grouped sequences.

## Method

**Participants.** Fifty-two members of the campus community at the University of York took part in the experiment in exchange for course credit or payment of £3 (approximately $5).

**Design and procedure.** The experiment manipulated a single independent variable: sequence type (ungrouped vs. grouped), which was a between-participants factor. Half the participants received the ungrouped sequences, whereas the remaining half received the grouped sequences.

The procedure was identical to that of Experiment 1, with the following exceptions. First, the sequence length was increased from seven to nine locations. By implication, intrusion errors were no longer possible, meaning that the likelihood that a particular error would be a transposition was increased relative to Experiments 1 and 2. Second, following the "begin trial" icon, all nine locations appeared simultaneously on-screen. After a 2,000-ms delay, the locations temporarily changed color from gray to yellow—one location at a time—according to a random sequence determined by the computer program controlling stimulus presentation. Each location was highlighted yellow for 500 ms, followed by a 250-ms interstimulus interval during which all

locations remained gray. Following the change in color of the final item, the locations disappeared for 1,000 ms, after which they reappeared prompting recall of the sequence. Third, for grouped sequences the interstimulus intervals separating the third and fourth and the sixth and seventh locations were 1,250 ms, creating the impression of three groups of three locations.

Participants were instructed to encode sequences visually, in the absence of supplementary verbal or gestural encoding strategies, and all participants once again reported compliance with these instructions. Each sequence type involved two practice trials followed by 80 experimental trials. Sessions lasted approximately 40 min.

## Results

**Accuracy serial position curves.** The accuracy serial position curves can be inspected in Figure 8A. As in Experiment 1, grouping enhanced overall recall accuracy and caused a scalloping of the

serial position curve, with effects of primacy and recency apparent within groups, as well as the sequence overall. Statistical confirmation of these trends was obtained by a 2 (sequence type) × 9 (serial position) ANOVA, which revealed significant main effects of sequence type, $F(1, 50) = 8.8$, $p < .01$, $r = .51$, and serial position, $F(8, 400) = 81.5$, $p < .001$, in conjunction with a significant interaction, $F(8, 400) = 9.23$, $p < .001$.

**Transposition gradients.** Figure 8B shows the transposition gradients, which exhibit the same features witnessed in the previous experiments. Of primary interest is whether grouping increased interpositions in grouped sequences. Such an outcome would be reflected by discontinuities in the transposition gradient for grouped sequences, with local peaks at transposition displacements ±3 and ±6. It is readily apparent that such peaks are absent in the data. As in Experiment 1, transpositions were probed further by classifying them according to whether they occurred within or
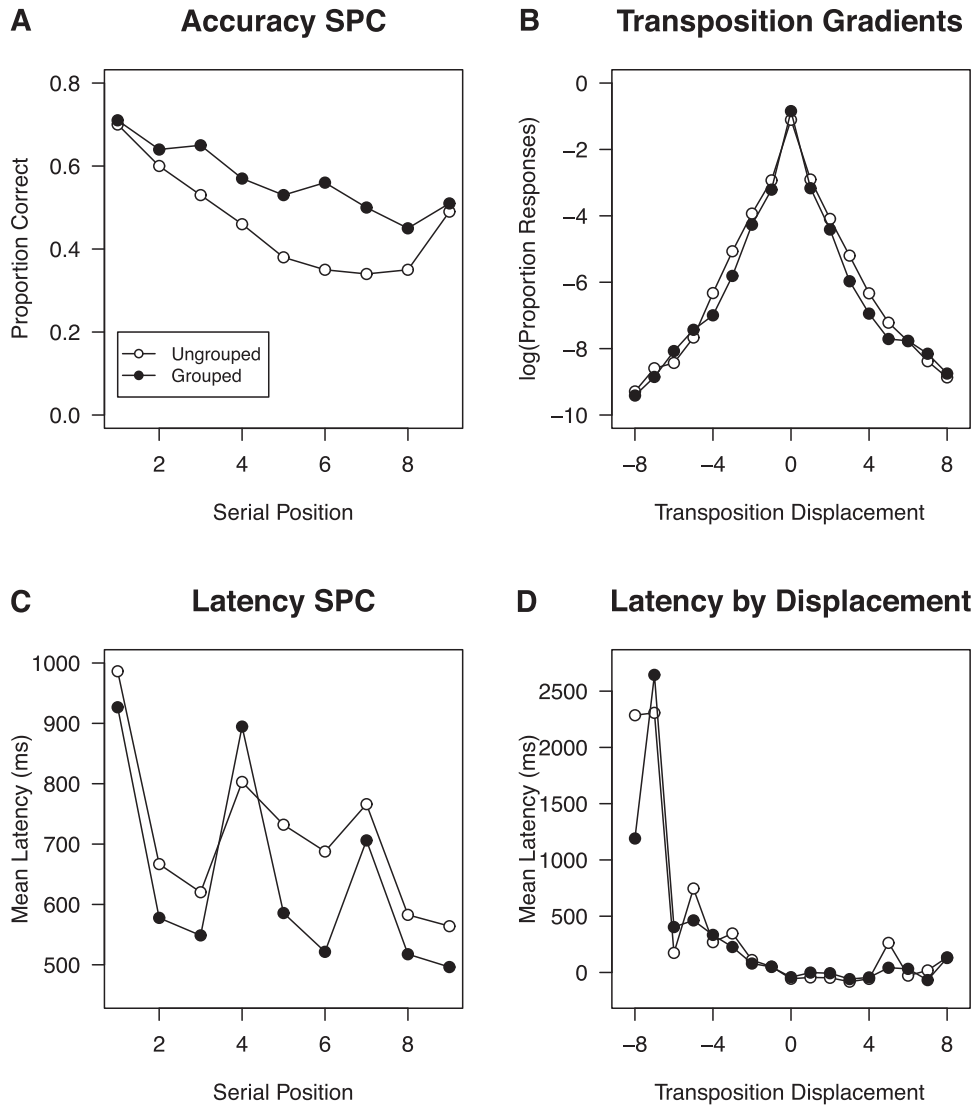


*Figure 8.* Response probability and recall latency data for Experiment 3. Panels show accuracy serial position curves (A), transposition gradients (B), latency serial position curves (C), and latency-displacement functions (D).

between groups, with the latter errors being further subdivided into interpositions and position nonpreserving errors. Paired comparisons performed on the log-odds transformed error proportions revealed that grouping increased the probability of transpositions within groups (ungrouped $M = .56$ vs. grouped $M = .70$), $t(50) = -4.35$, $p < .001$, $r = .52$, but decreased the probability of position nonpreserving transpositions between groups (ungrouped $M = .31$ vs. grouped $M = .19$), $t(50) = 5.45$, $p < .001$, $r = .61$. Crucially, grouping decreased—rather than increased—the probability of interpositions (ungrouped $M = .13$ vs. grouped $M = .11$), $t(50) = 2.09$, $p < .05$, $r = .28$.

**Latency serial position curves.** The latency serial position curves are plotted in Figure 8C, from which it is evident that the curve for grouped sequences exhibits pronounced peaks at Serial Positions 1, 4, and 7, corresponding to the beginning of groups, with a speed-up in the recall times at subsequent within-group positions. The curve for ungrouped sequences departs somewhat from what would normally be expected. Instead of following an inverted U-shape after the long initial output time, the curve follows a profile similar to that observed for grouped sequences, except that the peaks at Serial Positions 4 and 7 are less punctuated. One interpretation of this empirical pattern is that participants may have spontaneously grouped the ostensibly ungrouped sequences into threes during encoding. However, given the absence of any scalloping in the aggregate accuracy serial position curve—nor any systematic scalloping in the accuracy curves for individual participants—the small punctuations in the ungrouped latency curve may be better explained in terms of output buffering rather than spontaneous grouping. To explain, on the output buffering account, constraints on the number of motor responses that can be executed within close succession may have forced people to group their responses at output (see Farrell & Lelièvre, 2012, for further discussion).

Corroborating the aforementioned trends, a 2 (sequence type) × 9 (serial position) ANOVA performed on the mean recall latencies for correct responses revealed a significant main effect of serial position, $F(8, 400) = 31.15$, $p < .001$, but no significant main effect of sequence type, $F(1, 50) = 1.9$, $p = .174$, $r = .19$, and no significant interaction, $F(8, 400) = 1.97$, $p = .129$.

**Latency-displacement functions.** The filtered LDFs shown in Figure 8D confirm the empirical pattern observed in the previous two experiments. The overall trend is once again a negative one, but with a flattening of the slope of the function for postponements compared with anticipations: The recall times for anticipations increase the further an item is displaced, whereas with one unique deviation—the

single peak at displacement $+5$ for ungrouped sequences—the recall times for postponements are invariant with respect to the size of the displacement. The regression statistics for the LDF slope analyses are summarized in Table 3. They confirm that the slopes of the functions for anticipations were statistically negative: $t(25) = -4.11$, $p < .001$, $r = .63$, for ungrouped sequences, and $t(25) = -4.12$, $p < .001$, $r = .64$, for grouped sequences, whereas the slopes of the functions for postponements were weakly positive, but did not deviate significantly from zero: $t(25) = 1.55$, $p = .134$, $r = .30$, for ungrouped sequences, and $t(25) = .84$, $p = .41$, $r = .17$, for grouped sequences. Figure 9 shows the LDF slopes for individual participants. As in the previous experiments, the majority of participants contributed steep negative slope estimates for anticipations, whereas for postponements, there was an approximately equal distribution of negative and positive slopes, and, without exception, these were all shallow slopes.

## Discussion

As in Experiment 1, temporal grouping exerted a multiplicity of effects on recall accuracy, latency, and errors. Specifically, grouping caused an enhancement in overall recall accuracy, mini within-group primacy and recency effects, long output times prior to the production of the first item of each group, and a reduction in the frequency of transpositions between groups. Nevertheless, in line with Experiment 1, grouping did not engender an increase in the probability of interpositions. Turning to the LDFs, the empirical pattern observed is consistent with that witnessed in the previous two experiments. For both ungrouped and grouped sequences, the slopes of the functions for anticipations were negative, whereas the slopes of the functions for postponements were flat, conferring further support for a representational mechanism integrating a primacy gradient, position marking, and response suppression. Thus, consistent with the results of Experiment 1, the effects of temporal grouping point to a pivotal role for position marking in the representation of serial order in spatial sequences but the LDFs—as well as confirming the role of position marking—also identify a necessary role for a primacy gradient and response suppression.

The chief contribution of the present experiment has been to demonstrate that the empirical pattern of the LDFs observed in the previous experiments generalizes to the use of longer spatial sequences, a different temporal grouping pattern, and the use of a simultaneous—as opposed to sequential—spatial presentation array. The use of longer sequences is particularly diagnostic because it permitted an assessment of the relationship between recall latency and transposition displacement when transpositions could span a greater range of displacements, in order to determine whether the empirical pattern hitherto observed is subject to potential range effects. That the form of the LDF remains unaltered—despite the foregoing changes—is particularly telling, and further underscores the generality of this empirical signature of the primacy gradient, position marking, and response suppression mechanism.

## Quantitative Model Comparisons

Although the LDFs witnessed across the three experiments are most compatible with the theoretical prediction of the PG + PM + RS model, the possibility cannot be excluded that the models might predict qualitatively different LDFs when model parameters

Table 3

*Mean Regression Parameter Estimates for the Slopes of the Latency-Displacement Functions for Anticipations and Postponements in Experiment 3*

| Parameter | Estimate | SE | t | p | r |
|---|---|---|---|---|---|
| Ungrouped | | | | | |
| Anticipation | −213.48 | 52.00 | −4.11 | .00 | .63 |
| Postponement | 19.41 | 12.53 | 1.55 | .13 | .30 |
| Grouped | | | | | |
| Anticipation | −148.16 | 35.97 | −4.12 | .00 | .64 |
| Postponement | 4.66 | 5.56 | .84 | .41 | .17 |
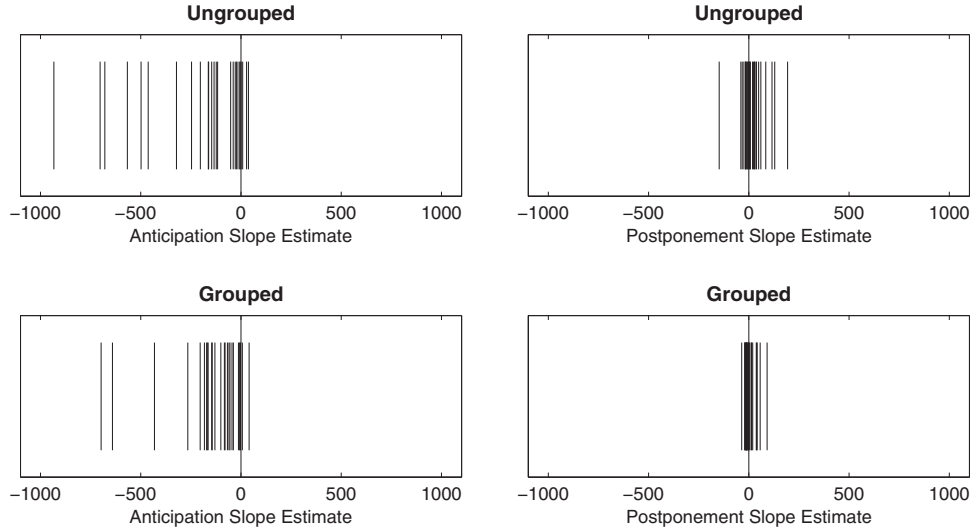
*Figure 9.* Individual regression parameter estimates for the slopes of the latency-displacement functions (LDFs) for anticipations and postponements in Experiment 3. Each vertical line in a panel represents a regression slope estimate for a single participant, for a particular experimental condition. The left-hand panels show the estimates for the slopes of the LDFs for anticipations; the right-hand panels show the estimates for the slopes of the LDFs for postponements. The first row of panels gives the slope estimates for ungrouped sequences; the second row of panels gives the slope estimates for grouped sequences. The full vertical line within each panel indicates a slope value of zero.

are estimated from the behavioral data. To establish whether this is the case, the five models were fit to the response probability and recall latency data for the ungrouped condition of Experiment 3. Of theoretical interest is whether models other than the PG + PM + RS model can reproduce the observed LDF when model parameter values are allowed to vary for each individual participant.

## Model Fitting

The five models were fit jointly to the accuracy and latency serial position curves, transposition gradients, and the LDFs (with the effect of output position removed) of individual participants using maximum likelihood parameter estimation (assuming normally distributed data with constant variance). The observed and predicted mean latencies for the latency serial position curves and LDFs were divided by $10^4$ to ensure that the variability in the response probability and latency measures were within a similar range so that they received equal weighting during the fitting process. The modeling procedure was the same as that employed for the initial simulations except that model parameter values were varied systematically using the SIMPLEX algorithm (Nelder & Mead, 1965) in order to maximize the log likelihood:

$$\ln L = \frac{-n}{2} \ln \left( \frac{RSS}{n} \right) \qquad (5)$$

where "ln" is the natural logarithm, *RSS* is the residual sum of squares, and *n* is the number of observations entering into the log-likelihood calculation (larger values of ln *L* indicate a better fit). Each parameter vector explored by the search algorithm involved 10,000 model simulation trials of nine-item sequences.

The parameters that were free to vary for the PM model were the weighting ($\lambda$) and distinctiveness ($\theta$) of the position markers.

These free parameters were augmented by the amount of response suppression ($\alpha$) in the PM + RS model and the amount of output interference ($\delta$) in the PM + OI model. The PG + RS model took as its free parameters the starting point ($a_1$) and steepness ($\gamma$) of the primacy gradient, and the degree of response suppression ($\alpha$). Finally, the free parameters for the PG + PM + RS model were the steepness of the primacy gradient ($\gamma$), the distinctiveness of the position markers ($\theta$), the weighting of the primacy gradient and position markers ($\omega$), and the degree of response suppression ($\alpha$). The two remaining parameters of this model (viz., $\lambda$ and $a_1$) were frozen to values of 1—which neutralizes any influence they might have—to minimize the number of to-be-estimated parameters. In addition to the aforementioned parameters, the iteration-to-ms scaling parameter *S* was included as a free parameter in all models. In summary, the number of free parameters was three for the PM model; four for the PM + RS, PM + OI, and PG + RS models; and five for the PG + PM + RS model.

The models were initially fit to the data according to the procedure just described, which yielded for each participant and each model, a set of best fitting parameter values, and an associated maximum log-likelihood estimate. However, as the models differ in their number of free parameters, it is necessary to augment this goodness-of-fit metric with a penalty term that punishes excessive model complexity. This is because it is well known that a more complex model with more free parameters can provide a better fit than a less complex model with fewer parameters (Pitt & Myung, 2002). Accordingly, in order to provide a measure of the descriptive accuracy of the models that takes into consideration differences in their degree of complexity, the log-likelihood estimates were converted into Akaike information criterion (AIC; Akaike, 1974) and Bayesian information criterion (BIC; Schwartz, 1978) scores. The AIC was calculated as

$$AIC_i = -2 \ln L_i + 2 V_i \quad (6)$$

where $V$ is the number of free parameters involved in maximizing $\ln L$, and $i$ indexes the model for which AIC is being calculated (smaller values of AIC indicate a better fit). As can be seen from Equation 6, the AIC rewards a model for its goodness of fit via its maximized log likelihood and punishes it as a function of its number of free parameters. Similarly, the BIC was calculated as

$$BIC_i = -2 \ln L_i + V_i \ln(n) \quad (7)$$

Like the AIC, the BIC rewards a model for its goodness of fit via its maximized log likelihood, but punishes it as a function of the number of free parameters weighted by the number of observations entering into the log-likelihood calculation (smaller values of BIC indicate a better fit). Accordingly, the BIC offers a more stringent correction for model complexity.

Because AIC and BIC scores are discrete rather than continuous values, it can often be difficult to discern whether differences in AIC and BIC between models are trivial or meaningful. Accordingly, to facilitate interpretation of the model comparisons, the raw AIC and BIC scores were converted into so-called IC weights (Burnham & Anderson, 2002; Lewandowsky & Farrell, 2011; Wagenmakers & Farrell, 2004), which express the evidence in favor of each model on a continuous measure of evidence. The IC weight for model $i$ was calculated by

$$wIC_i = \frac{\exp(-0.5 \, \Delta IC_i)}{\sum_{k=1}^{K} \exp(-0.5 \, \Delta IC_k)}, \quad (8)$$

where $\Delta IC_i$ is the difference in IC between model $i$ relative to the best model, and each $\Delta IC_k$ is the difference in IC between a specific model $k$ in the candidate set $K$ and the best model. These IC weights—normalized to sum to 1—represent the probability that each model is the best given the data and the competitor models under comparison. The support for a model is considered equivocal if its IC weight does not exceed $1/N$, where $N$ is the number of models under comparison. Thus, with five models, the support for a particular model is considered equivocal if its IC weight does not exceed 0.2.

Table 4

*Parameter Estimates for the Fits of the Models to the Ungrouped Condition of Experiment 3 Averaged Across Participants*

| Model | λ | θ | $a_1$ | γ | ω | α | δ | S |
|---|---|---|---|---|---|---|---|---|
| PM | 0.66 | 0.66 | — | — | — | — | — | 27.08 |
| PM + RS | 0.62 | 0.62 | — | — | — | 0.98 | — | 28.30 |
| PM + OI | 0.70 | 0.58 | — | — | — | — | 0.44 | 34.75 |
| PG + RS | — | — | 0.69 | 0.88 | — | 0.98 | — | 29.33 |
| PG + PM + RS | — | 0.77 | — | 0.88 | 0.52 | 0.97 | — | 32.31 |

*Note.* λ = weighting of position markers; θ = distinctiveness of position markers; $a_1$ = starting point of primacy gradient; γ = steepness of primacy gradient; ω = weighting of primacy gradient and position markers; α = degree of response suppression; δ = amount of output interference; S = iteration-to-ms scaling; PM = position marking; RS = response suppression; OI = output interference; PG = primacy gradient.

## Model Selection

The average parameter estimates for the five models can be inspected in Table 4, and the IC weights and associated goodness-of-fit quantities can be scrutinized in Table 5. It can be seen from inspection of the latter table that the average AIC weight for the PG + PM + RS model is 0.66, which is reliably different from 0.2, $t(25) = 6.86$, $p < .001$, $r = .81$, based on a one-sample $t$ test, and distinctly larger than the average weights of the other four models. The PM + OI model yielded the next largest average weight, followed by the PG + RS, PM, and PM + RS models. The BIC weights also suggest that the PG + PM + RS model is the preferred model. The average BIC weight for this model is 0.56, which is reliably different from 0.2, $t(25) = 4.84$, $p < .001$, $r = .7$, and sufficiently larger than the BIC weights for the other four models. As for the AIC weight comparisons, the PM + OI model yielded the next largest average weight, followed by the PG + RS, PM, and PM + RS models.[4] That the PM + OI model provided a better fit than the PG + RS model is surprising in light of the fact that the qualitative LDF predicted by the latter model at the outset is a better approximation of the empirical LDF. Although the PG + RS model still predicted a more realistic LDF than the PM + OI model, the latter model provided a better description of the serial position curves and transposition gradients, which accounted for its slightly better overall fit.

## Model Predictions

The predictions of the models are portrayed in Figure 10, from which it can be seen that the accuracy serial position curves predicted by the PM + RS, PM + OI, PG + RS, and PG + PM + RS models (Figure 10A) show good correspondence with the empirical data (Figure 8A). Notably, all four models accommodate the more extensive primacy than recency witnessed empirically. By contrast, the PM model predicts a symmetrically bowed serial position curve that is contrary to the empirical data. The quality of agreement between the predicted and observed transposition gradients (Figure 10B vs. Figure 8B) is also good. In accordance with the empirical data, all models predict steeply peaked transposition gradients in which adjacent-neighbor transpositions predominate, and the error gradients for anticipations and postponements are approximately symmetrical. However, the predicted latency serial position curves (Figure 10C) are at variance with the empirical data (Figure 8C). Whereas the observed latency curve peaks at the first output position, with less punctuated peaks at Serial Positions 4 and 7, the predicted latency curves for all models follow an inverted-U-shaped profile and are devoid of the discontinuities witnessed in the data. Consistent with our earlier simulations, it is apparent that the five models cannot be qualitatively differentiated

---

[4] Astute readers will have noticed that in Table 5, the AIC and BIC values for the PM + RS model are smaller than for the PM model—indicating that the former model provided a better fit than the latter—but paradoxically, the average AIC and BIC weights are slightly larger for the PM model. This situation materialized because for one individual participant, the AIC and BIC weights of the PM model were markedly larger than the weights for other participants fitted by this model, which artificially elevated the average AIC and BIC weights of this model above those of the PM + RS model (the weights are also very small numbers, which renders them vulnerable to such an effect).

Table 5
*AIC and BIC Weights With Associated Goodness-of-Fit Quantities for the Fits of the Models to the Ungrouped Condition of Experiment 3 Averaged Across Participants*

| Model | k | ln L | AIC | ΔAIC | wAIC | BIC | ΔBIC | wBIC |
|---|---|---|---|---|---|---|---|---|
| PM | 3 | 135.62 | −265.25 | 44.54 | 0.0138 | −259.61 | 41.57 | 0.0251 |
| PM + RS | 4 | 144.71 | −281.42 | 28.37 | 0.0071 | −273.91 | 27.28 | 0.0110 |
| PM + OI | 4 | 153.57 | −299.14 | 10.65 | 0.1945 | −291.64 | 9.54 | 0.2321 |
| PG + RS | 4 | 151.97 | −295.93 | 13.86 | 0.1213 | −288.51 | 12.68 | 0.1718 |
| PG + PM + RS | 5 | 159.19 | −308.38 | 1.40 | **0.6633** | −299.03 | 2.16 | **0.5599** |

*Note.* $k$ = number of free model parameters; ln $L$ = log maximum likelihood; AIC = Akaike information criterion; ΔAIC = difference in AIC with respect to the best fitting model; wAIC = AIC weight (with the best model's weight in bold face); BIC = Bayesian information criterion; ΔBIC = difference in BIC with respect to the best fitting model; wBIC = BIC weight (with the best model's weight in bold face); PM = position marking; RS = response suppression; OI = output interference; PG = primacy gradient.

on the basis of their predicted serial positions curves and transposition gradients.

To adjudicate between the models, we now turn to their predicted LDFs, which are illustrated in Figure 10D with the effects of output position removed, as per the data. To facilitate comparison, the LDFs in this figure have been centered at displacement 0 by subtracting the predicted mean latencies of the PM, PM + RS, PM + OI, and PG + PM + RS models at this displacement from the predicted mean latency of the PG + RS model—which had the fastest mean reaction time at this displacement—before adding the resulting values (minimum = 40.33 ms; maximum = 115.02 ms) to the predicted mean latencies of the models for all displacement values. Consistent with the initial predictions (Figure 4D), it can be seen that all models predict a negative anticipation slope, but the model predictions for postponements differ widely. As before, the PM, PM + RS, and PM + OI models all predict steep positive postponement slopes; the PG + RS model predicts a negative postponement slope—albeit a somewhat shallower slope compared with the initial qualitative prediction of this model—whereas the PG + PM + RS model predicts a flat postponement slope. Inspection of the data in Figure 8D lends empirical support for the PG + PM + RS model—the slope of the LDF for postponements for ungrouped sequences is approximately flat. Thus, the current simulations demonstrate that even when the models are fit directly to representative empirical data, only the PG + PM + RS model is able to reproduce the error latency profile observed empirically.

Although the model comparisons clearly identify the PG + PM + RS model as the preferred model of the data, one potential stumbling block in interpreting the results of the simulations is that the latency curves predicted by the models depart markedly from the empirical data. This raises the question of whether the models might predict qualitatively different LDFs if they were able to accommodate this aspect of the data. Allaying such concerns, we report further simulations in the Appendix, which show that, under such conditions, the LDFs predicted by the models remain qualitatively unaltered from those presented here.

## Parameter Sensitivity Analysis

The results of the model fitting exercise confirm that only the PG + PM + RS model—of the competitor models under comparison—can reproduce the observed LDF when the models are fit directly to the behavioral data. To probe model behavior more

deeply, we next examine the sensitivity of the models' predicted LDFs to variations in their parameter settings—a technique dubbed "parameter sensitivity analysis" (e.g., Li, Lewandowsky, & DeBrunner, 1996). A parameter sensitivity analysis can establish whether a model's behavior stems from its core representational assumptions. If so, the model's behavior should remain stable across reasonable variations of its parameter values. By contrast, if the behavior is restricted only to a narrow range of parameter settings, then it may result from arbitrary and unprincipled properties of the model.

Farrell and Lewandowsky (2004) conducted a parameter sensitivity analysis of the PM, PM + RS, PM + OI, and PG + RS models. For each model, they specified a broad range of values for each of its parameters before factorially combining these to create a grid of parameter setting combinations. Predictions were then obtained for each model, for each point in its grid space. This yielded for each model a large number of LDFs covering a wide range of parameter settings. The dependent measure was the slope of the LDFs for postponements only. Their simulations revealed that the models employing position marking to represent serial order—either in isolation (PM) or in conjunction with response suppression (PM + RS) or output interference (PM + OI)—consistently predicted positive postponement slopes, whereas the PG + RS model consistently predicted negative postponement slopes. The simulations therefore confirmed that the predictions of these four models scrutinized thus far are representative of their general behavior. However, crucially, Farrell and Lewandowsky (2004) did not include the PG + PM + RS model in their sensitivity analysis because this model was not considered until a later article (Lewandowsky & Farrell, 2008), in which the authors presented only qualitative predictions of the model.

Accordingly, to examine the consistency of the predictions of the PG + PM + RS model, we replicated the sensitivity analysis of Farrell and Lewandowsky (2004), but this time incorporated the PG + PM + RS model into the model comparisons. The same parameters used to fit each model to the ungrouped condition of Experiment 3—excluding the iteration-to-ms scaling parameter $S$, which was fixed to a value of 1—were varied independently from 0.05 to .95 in steps of 0.1. This meant that two parameters were varied for the PM model; three for the PM + RS, PM + OI, and PG + RS models; and four for the PG + PM + RS model. The parameter values were factorially combined to create a grid of parameter-setting combinations. This yielded 100 parameter-
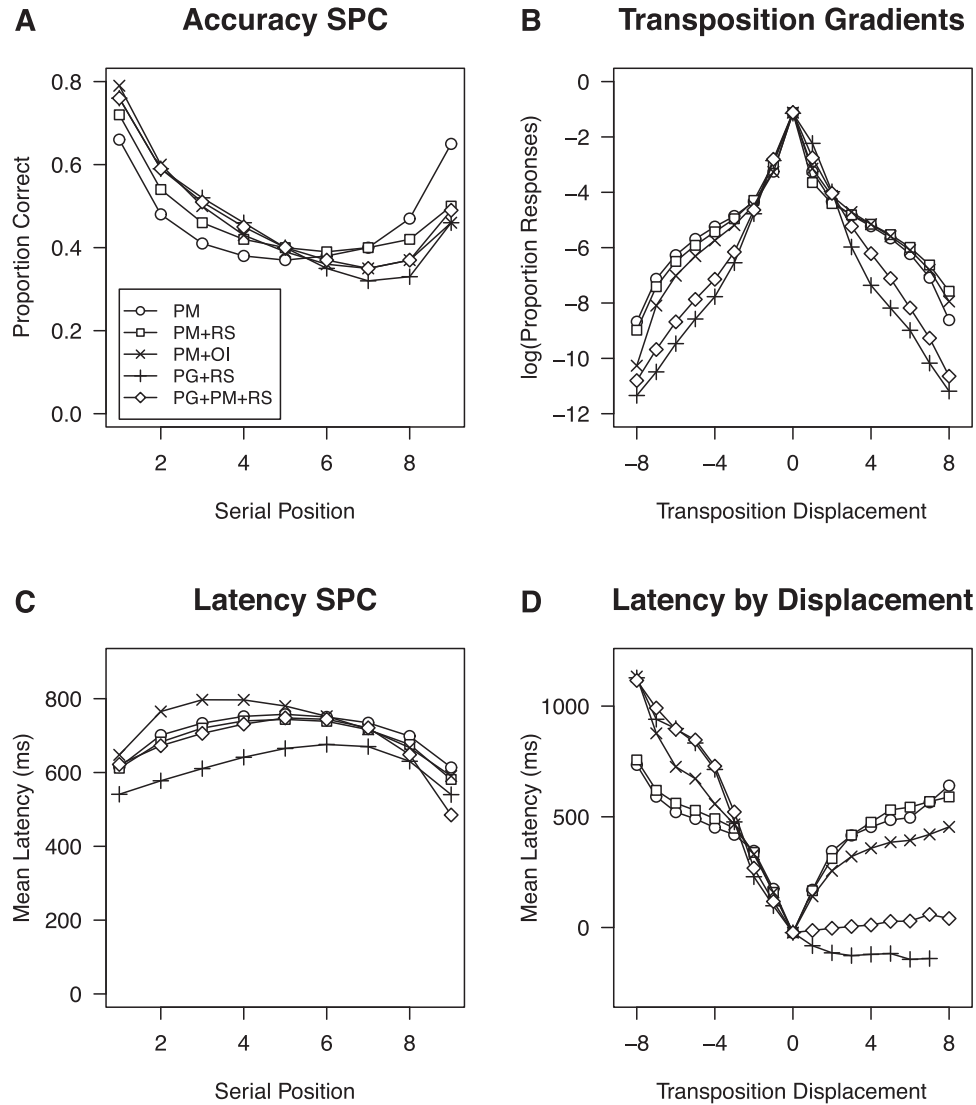
*Figure 10.* Fits of five models of serial order to the ungrouped condition in Experiment 3. Panels show accuracy serial position curves (A), transposition gradients (B), latency serial position curves (C), and latency-displacement functions (D). PM = position marking; RS = response suppression; OI = output interference; PG = primacy gradient.

setting combinations for the PM model ($10^2$); 1,000 combinations for the PM + RS, PM + OI, and PG + RS models ($10^3$); and 100,000 combinations for the PG + PM + RS model ($10^4$). All remaining parameters were set to the same values used to generate the initial model predictions in Figure 4. For each combination of parameter values, 1,000 simulation trials were run of nine-item sequences.

To facilitate comparison of the model predictions with the data, the individual postponement slopes predicted by each of the models were first converted from measurement in model iterative cycles to milliseconds. This was accomplished using the following three-stage procedure. In the first stage, regression analyses were performed for each model, in which the LDF predicted under each combination of parameter settings (with the effects of output position subtracted) was entered as a predictor variable and the

aggregate LDF for ungrouped sequences in Experiment 3 was used as the dependent variable. For each predicted model LDF, this yielded two scaling parameters: an intercept parameter (in milliseconds) and an iteration-to-millisecond slope parameter. In the second stage, each predicted model LDF was transformed from iterations to milliseconds employing the scaling parameters obtained from the first-stage analyses. In the final stage, another set of regression analyses were performed in order to obtain the postponement slope estimates for the transformed LDFs of each of the models.

Before turning to the outcomes of the simulations, we first consider the pattern of results expected from the PG + PM + RS model if it is a preferable model of the data. Although the aggregate LDFs observed empirically and the aggregate LDF predicted by this model are approximately flat, in both cases, the individual

LDFs from which they are composed consist of a combination of shallow negative and positive postponement slopes. Thus, the approximately flat aggregate postponement slopes for both data and model actually represent the result of averaging over a mixture of shallow negative and positive postponement slopes. This pattern is illustrated graphically for the data in Figure 11A, which is a density histogram showing the distribution of individual participant postponement slopes for the various conditions of the three experiments. It can be seen by inspection that the majority of observed postponement slopes (92%) are shallow positive or negative slopes falling in the range of −100 ms to 100 ms. Accordingly, we would expect the PG + PM + RS model to predict a majority of shallow negative and positive postponement slopes falling within this empirical range as its main theoretical prediction. More generally, the distribution of postponement slopes predicted by this model should correspond closely with the empirical distribution shown in Figure 11A.

The predictions of the models are shown alongside the data in Figure 11. Replicating the sensitivity analysis of Farrell and Lewandowsky (2004), the PM, PM + RS, and PM + OI models (Figure 11B, C, and D, respectively) predicted a majority of positive postponement slopes (84%, 88%, and 87% of model predictions, respectively), whereas the PG + RS model (Figure 11E) predicted a majority of shallow negative postponement slopes (89% of model predictions), confirming that the qualitative predictions of the models observed thus far represent core predictions of the models. Critically, the PG + PM + RS model (Figure 11F) predicted a majority of shallow postponement slopes: 91% of

the model's predictions fell within the range of −100 ms to 100 ms. Indeed, the distribution of slopes predicted by the model is strikingly similar to the distribution of slopes observed empirically (Figure 11A). The results of the sensitivity analysis thus confirm that the LDF predicted by the PG + PM + RS model under its best-fitting parameter values follows from its core representational assumptions.

## General Discussion

The three experiments reported in this article examined the dynamics of transpositions in a spatial serial recall task in order to test the error latency predictions of five alternative mechanisms for the representation of serial order. The results of the experiments consistently revealed that transposition latency is a negative function of transposition displacement, but with a reduction in the slope of the function for postponements compared with anticipations. This empirical pattern is uniquely predicted by a competitive queuing mechanism within which serial order is represented by combination of a primacy gradient of activations over items and associations between items and position markers, with suppression of items following recall. The same empirical pattern is incompatible with the four alternative mechanisms for representing serial order. These results are consistent with those reported by Farrell and Lewandowsky (Farrell & Lewandowsky, 2004; Lewandowsky & Farrell, 2008) for the serial recall of verbal sequences, and provide the first clear evidence that spatial and verbal STM rely on common principles for the representation of serial order.
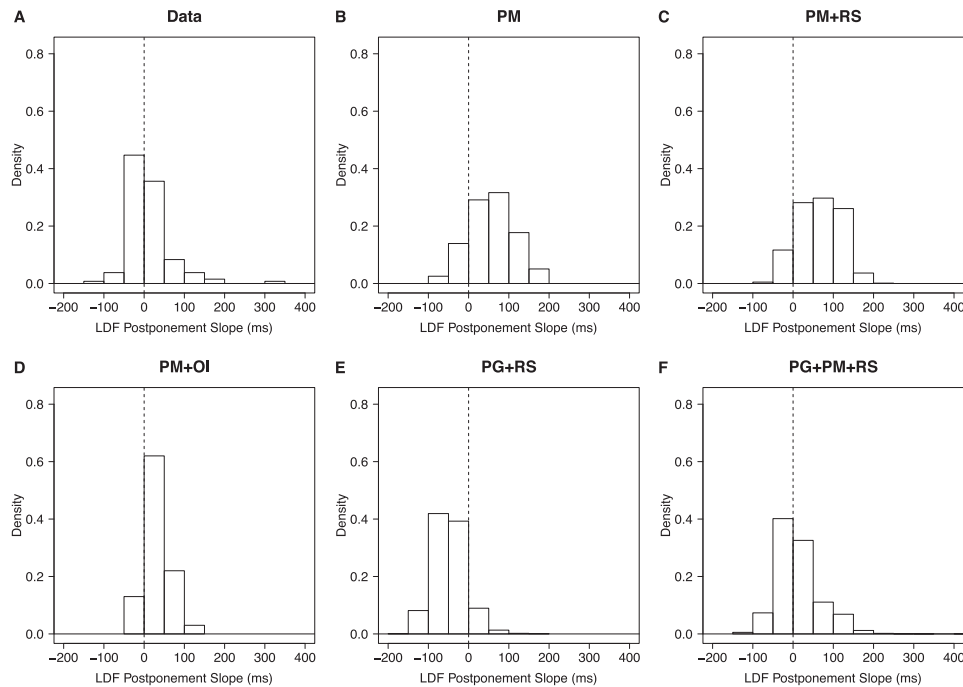


*Figure 11.* Distributions of latency-displacement function (LDF) postponement slopes across individual participants and the parameter space of the models. Panels show the slope distributions for the data (A), PM model (B), PM + RS model (C), PM + OI model (D), PG + RS model (E), and PG + PM + RS model (F). The broken vertical line in each panel corresponds to a slope value of 0. PM = position marking; RS = response suppression; OI = output interference; PG = primacy gradient.

The generality of the results of the current experimental series is indicated by the fact that LDFs exhibiting flat postponement slopes were observed across manipulations of temporal grouping (Experiments 1 and 3), postsequence interference (Experiment 2), sequence length, and presentation format (Experiment 1 and 2 vs. Experiment 3). We note also that LDFs characterized by flat postponement slopes are a feature of an unpublished experiment reported by Hurlstone (2010, Experiment 12) who, instead of employing a fixed set of spatial locations (as in the current experiments), employed spatial locations whose coordinates varied randomly from trial to trial. The consistency of this empirical outcome suggests that a primacy gradient, position marking, and response suppression are core principles of serial order in spatial STM. Qualified support for the role of the three representational principles was provided by the results of the quantitative-model-fitting exercise, which confirmed that only the PG + PM + RS model could reproduce the observed LDFs when model parameters were estimated directly from the behavioral data. The predictions of the model were also shown to be robust to broad variations of its parameter settings, indicating that the model's behavior follows from its core representational principles.

## Potential Limitations

Before embarking on a discussion of the theoretical implications of our results, we digress briefly by considering a potential limitation of our modeling approach—and, by extension, the approach of Farrell and Lewandowsky (2004), upon which it is based. To keep our simulations simple, we employed a single-layer lateral inhibition network to test the recall latency predictions of the different seriating mechanisms. However, the competitive queuing architecture consists of two layers—a parallel planning layer and a competitive choice layer—and in a fully implemented connectionist instantiation of the competitive queuing model (e.g., Bullock & Rhodes, 2003; Davelaar, 2007; Grossberg & Pearson, 2008), the activations in both layers update dynamically over time. Crucially, however, the dynamics of the two layers are different. Because the purpose of the parallel planning layer is to preserve the serial plan for the sequence, items in this layer compete only weakly with one another via lateral inhibition, which causes the activations in this layer to evolve relatively slowly over time. By contrast, because the purpose of the competitive choice layer is to select a single representation from among a set of parallel activated representations, item representations in this layer compete strongly via lateral inhibition, which causes the activations to evolve relatively rapidly. The different dynamics in the two layers means that they will interact nonlinearly in a manner that is not captured by the current single-layer architecture. It therefore follows that the deliberately simplified seriating mechanisms examined here might predict somewhat different error latency profiles in a fully implemented two-layer connectionist competitive queuing architecture. We defer an assessment of this possibility for another occasion, but note, for now, that we are confident that the predictions of the mechanisms will generalize under such circumstances.

## Complementary Evidence for Position Marking

The temporal grouping manipulation employed in Experiments 1 and 3 was incorporated to provide further information about the role of position marking in spatial STM. Consistent with a wealth of verbal studies (Farrell & Lelièvre, 2009; Farrell & Lewandowsky, 2004; Frankish, 1985, 1989; Henson, 1999; Maybery et al., 2002; Parmentier & Maybery, 2008; Ryan, 1969a, 1969b), as well as the single spatial study of Parmentier and colleagues (2006, Experiment 4), temporal grouping exerted a multiplicity of effects on recall performance. Specifically, grouping enhanced the accuracy of recall, and caused effects of primacy and recency within groups, long output times prior to the production of the first item of each group, and a reduction in the frequency of transpositions between groups. These results provide independent and complementary evidence that positional information contributes to the encoding of spatial sequences.

Nevertheless, one prominent and pervasive feature of temporal grouping observed in studies of verbal STM that failed to manifest in the experiments reported here—and in the study of Parmentier et al. (2006)—is the increase in interpositions associated with temporally grouped sequences (Farrell & Lelièvre, 2009; Farrell & Lewandowsky, 2004; Henson, 1996, 1999; Ng & Maybery, 2002, 2005; Ryan, 1969a, 1969b). In the verbal domain, this observation constitutes a key piece of empirical support for positional theories of serial recall. Accordingly, it is important to ascertain why this effect of grouping does not generalize to the spatial domain. In attempting to interpret this discrepancy, we consider three possible explanations.

The first possibility is that there may be methodological discrepancies between the way in which grouping effects were elicited in our own experiments and in the verbal studies that may have obviated against the observation of interpositions. This explanation seems unlikely, given that we employed patterns of grouping and temporal presentation schedules akin to those employed in verbal studies and were successfully able to elicit the other major empirical referents of grouping with our manipulations. Moreover, in an as-yet-unpublished study, Hurlstone (2010; Experiment 7) reports an experiment that directly compared grouping effects in spatial and verbal STM using the same methodology—closely matched to Experiment 3—in which kindred effects of grouping on recall accuracy, latency, and the pattern of within- and between-group transpositions were observed in grouped spatial and verbal serial recall, but an increase in interpositions was witnessed only for grouped verbal serial recall.

A second possibility is that our failure to observe interpositions in grouped spatial serial recall might be a consequence of the way our spatial sequences were constructed. Although temporal factors exert a strong effect on error production in spatial tasks—as indicated by the detrimental effect of sequence length on recall accuracy (Smyth, 1996; Smyth & Scholey, 1994), the locality constraint underlying transpositions (Parmentier et al., 2006; Smyth & Scholey, 1996), and the effects of temporal grouping (current Experiments 1 & 3; Hurlstone, 2010; Parmentier et al., 2006)—spatial constraints exert an effect also. Specifically, studies have shown that properties of spatial sequences—such as the relative distance between successive locations (Guérard, Tremblay, & Saint-Aubin, 2009; Parmentier et al., 2006), the number of crosses in the sequence path (Parmentier & Andrés, 2006; Parmentier et al., 2005), the exact configuration of locations (Kemps, 1999), and the extent to which they can be segregated into subgroups based on Gestalt organizational principles (De Lillo, 2004; De Lillo & Lesk, 2010; Kemps, 2001; Parmentier et al., 2005; Rossi-Arnaud, Pieroni, & Baddeley, 2006)—exert effects on error production. It seems reasonable to speculate that such spatial con-

straints, which were not controlled in our experiments, may have interacted in unanticipated ways with our temporal grouping manipulation, causing a shift in the expected pattern of recall errors.

The third possibility—and our preferred interpretation of the data—is that positional information in grouped sequences may be represented differently in the spatial and verbal domains. Specifically, whereas grouping effects in the verbal domain are best understood by recourse to a multidimensional representational scheme, within which one dimension represents the positions of groups in the sequence and a second represents the positions of items within groups, grouping effects in the spatial domain might be better understood by recourse to a multidimensional representational scheme, within which the second dimension represents information about the positions of items in the sequence overall, rather than within groups. Such a representational scheme would be sufficient to reproduce the effects of grouping documented in the current experiments, and, in line with the data, would not predict an elevation in interpositions in grouped sequences.

In summary, although we can reject the first possibility, to adjudicate between the second and third possibilities, further experiments will be required that systematically manipulate different properties of the spatial sequences to-be-recalled to establish whether interpositions materialize under some spatial stimulus conditions but not others, or whether the absence of such errors is a robust phenomenon that reflects a fundamental difference in the way positional information is represented in the spatial and verbal domains.

## Complementary Evidence for Response Suppression

To further test the involvement of response suppression in spatial STM, we inspected the incidence of erroneous repetitions across the three experiments. Repetitions are rare and widely separated in the serial recall of verbal sequences, accounting for approximately 2% (Henson, 1996) to 5% (Vousden & Brown, 1998) of all responses, with an average lag of three to four positions between the two occurrences of the repeated item (Henson, 1996). The scarcity of erroneous repetitions has been taken as evidence that, in verbal STM, items are suppressed or removed from memory once they have been recalled. That the two instances of the repeat tend to be well separated has further been taken as evidence that this response suppression gradually wears off over time (although see Duncan & Lewandowsky, 2005, for evidence to the contrary). Surprisingly, to our knowledge, no studies have yet reported repetition data for the serial recall of sequences of spatial locations. Indeed, in most studies, repetitions are not even possible, because once a location has been selected, its color is changed to indicate that it has been chosen, and selection of that location again is prohibited by the computer program controlling response collection. By contrast, in the serial recall task employed here, once a location was selected, its color only changed transitorily to indicate that the response had been registered, after which it could be chosen again, thereby permitting an examination of the incidence of erroneous repetitions.

In the current experiments, repetitions accounted for approximately 1% of all responses—an occurrence rate well below that expected by chance. The two instances of the repeat were also widely separated: The average lag between the two occurrences of the repeat was four to five positions across the seven-item sequence conditions of Experiments 1 and 2, and six to seven positions across the nine-item

sequence conditions of Experiment 3. The paucity of erroneous repetitions in the ordered recall of spatial sequences, accompanied by their wide temporal separation, confers further support for the operation of response suppression in spatial STM.

## Implications for Accounts of Working Memory

We now attempt to situate the inferred principles within the broader theoretical framework of the working memory model of Baddeley and Hitch (Baddeley, 1986, 2000, 2007; Baddeley & Hitch, 1974). As noted in the introduction, according to the working memory model, STM consists of separate subsystems for the retention of verbal and visuospatial information—known as the phonological loop and visuospatial sketchpad, respectively. The latter subsystem is hypothesized to be further fractionated into separate visual and spatial subcomponents—dubbed by Logie (1995) as the "visual cache" and "inner scribe." The working memory model also postulates the existence of an attentional control system—known as the "central executive"—that is responsible for coordinating the working memory storage subsystems, and an "episodic buffer" whose function is to integrate—"bind together"—information held in the STM subsystems and long-term memory. As previously highlighted, although the working memory model has been successful in explaining a wealth of phenomena of STM, one widely acknowledged limitation is that it does not specify the detailed mechanisms by which serial order is processed.

In a recent review of the STM literature (Hurlstone et al., 2014), we proposed that the phonological loop and visuospatial sketchpad function as competitive queuing parallel sequence planning systems. This proposal is buttressed by the explanatory success of two explicit computational theories of the phonological loop—namely, the primacy model (Page & Norris, 1998) and the Burgess and Hitch (1999) model—both of which utilize the competitive queuing mechanism to generate serial order and are capable of simulating an impressive set of benchmark data on serial recall. Both models employ a primacy gradient of activations to represent serial order in conjunction with response suppression during recall output. The model of Burgess and Hitch (1999) additionally incorporates associations between items and a positional context signal to represent order among items. There is, by now, an impressive array of data—which includes the pattern of transposition error latencies observed by Farrell and Lewandowsky (2004)—conferring support for each of these representational principles in verbal STM, suggesting that they are core constructs of the phonological loop. In our review of the extant literature (Hurlstone et al., 2014), we were unable to specify the principles underlying the operation of the visuospatial sketchpad because of a shortage of direct empirical evidence. However, the current study goes some way toward plugging this theoretical gap.

Specifically, the results of the current empirical and modeling exercise suggest that a primacy gradient, position marking, and response suppression are also core constructs of the visuospatial sketchpad, at least the spatial component of this subsystem—namely, the "inner scribe"—that the Corsi-Blocks Task is taken to index (e.g., Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999). Readers may reasonably ask whether this renders the notion of functionally distinct spatial and verbal STM subsystems redundant. Can the data not be better and more parsimoniously explained by invoking a unitary memory system? Although some

theorists have indeed argued that similarities between spatial and verbal STM are suggestive of a unitary store (Jones, Beaman, & Macken, 1996; Jones et al., 1995), such a view is contradicted by a wealth of data demonstrating double dissociations between spatial and verbal STM at the behavioral, neuropsychological, and neuroimaging levels (see Baddeley, 2007, for a review). The current results in no way compromise the notion of distinct spatial and verbal STM stores for items: They merely suggest that the two subsystems process serial order in functionally similar ways.

Nevertheless, it is meaningful to ask whether, within the working memory architecture, the mechanisms that instantiate these principles across domains are entirely functionally divorced from one another. Starting with the primacy gradient, the simulations reported here assume that the primacy gradient is implemented over the representations of items in a STM store.[5] Thus, given the data pointing to the existence of distinct storage subsystems for spatial and verbal memoranda just alluded to, this would imply that separate primacy gradients represent serial order in the spatial and verbal STM subsystems. Turning to response suppression, one possibility is that a shared mechanism mediated by an executive control system—the central executive component of the working memory model—is responsible for inhibiting the representations of recalled items in the STM subsystems. This comes with the corollary that response suppression is, to a large extent, a strategic and willfully controlled act. The alternative possibility—and our preferred view—is that response suppression is implemented locally within the STM subsystems. This is more consistent with the conceptualization of response suppression in computational theories of serial recall, which envisage suppression as an obligatory process that is not under volitional control, and is supported by the finding of Henson (1998b) that participants are extremely poor at recalling the second instance of a repeated item even when those repetitions are detected with a very high level of accuracy (85%).

Different comments apply regarding the mechanism responsible for representing positional information in STM. Because the positional context signals employed in positional theories of serial recall (e.g., Brown et al., 2000; Burgess & Hitch, 1999; Henson, 1998a) represent serial order information independently of item information—unlike the primacy gradient mechanism, in which item and order information are conjunctively coded—it is possible that a common context signal, alienable from the STM subsystems themselves, might represent positional information across different STM domains. Indeed, Burgess and Hitch (1999) were explicit in suggesting that the positional context signal in their network model of the phonological loop might also be responsible for coding the position of nonverbal items. In terms of the working memory model, one speculation is that the positional context signal maps onto the episodic buffer (Baddeley, 2000; see Burgess & Hitch, 2005, for a similar suggestion). In keeping with the proposed binding function of the buffer, this would enable the same context signal to be flexibly associated with items from different modalities and domains. It would also provide a basis for encoding cross-domain sequences (e.g., a mixed sequence containing spatial and verbal items; cf. Farrell & Oberauer, 2014), because items emanating from different sources would be bound to the same context signal. Of course it is possible that the STM subsystems each possess their own dedicated positional context signals. Although less parsimonious, this possibility should not be dismissed especially in light of our failure to fully replicate the effects of

temporal grouping documented in verbal STM, which may allude to domain-specific context signals with subtly different properties.

One question left unaddressed by the present study is how serial order is represented within the visual component of the visuospatial STM subsystem—namely, the "visual cache." Empirically, visual STM for serial order has been examined by presenting people with sequences containing novel visual patterns—created by randomly filling the cells of visual matrices—or unfamiliar faces, conveyed from a constant spatial position. At recall, the items are simultaneously represented in a jumbled arrangement and the participant's task is to sort the items back into their original presentation order. Studies employing such stimuli have shown that visual STM for serial order exhibits serial position curves and transposition gradients akin to those witnessed in spatial and verbal STM (Avons, 1998; Avons & Mason, 1999; Smyth, Hay, Hitch, & Horton, 2005; Ward, Avons, & Melling, 2005), suggesting that principles of serial order in verbal STM are extensible not only to the spatial domain but also to the visual domain. However, direct evidence for those representational principles is currently lacking because—as we have shown here—these features of memory for serial order are explicable in terms of a variety of mechanisms for representing serial order. Accordingly, it will be fruitful to examine the pattern of transposition latencies accompanying performance on the visual STM tasks just described in order to establish which combination of the principles examined here is most likely to contribute to the representation of serial order in the visual domain, and whether those principles are the same as the ones identified in the current work.

## Closing Remarks

Before closing, we wish to underscore that we are not proposing that all phenomena of spatial and verbal STM for serial order can be explained by reference to a common set of explanatory constructs. In addition to the shared features in the data that mandate common principles and mechanisms, there also exist domain specific findings that any adequate models of spatial and verbal STM for serial order must be constrained by. Principal among these in the spatial domain are the results of studies demonstrating how Gestalt organizational principles of visual perception support the encoding of spatial sequences (De Lillo, 2004; De Lillo & Lesk, 2010; Kemps, 2001; Parmentier et al., 2005; Rossi-Arnaud et al., 2006); the results of studies illuminating the nature of the reference frames used to represent spatial locations (Avons, 2007; Avons & Oswald, 2008; Bernardis & Shallice, 2011); and the results of secondary-task interference studies showing how shifts of spatial selective attention and task-irrelevant eye movements impair memory for spatial sequences (Lawrence, Myerson, & Abrams, 2004; Pearson & Sahraie, 2003; Smyth, 1996; Smyth & Scholey, 1994; Tremblay, Saint-Aubin, & Jalbert, 2006). The common explanatory constructs identified by the current research will no doubt need to be augmented with bespoke mechanisms and as-

---

[5] Another way to implement the primacy gradient would be in the strength of the associations between items and their position markers. However, this approach is incompatible with the error latency data reported here. Specifically, we have shown by simulation (not reported) that a primacy gradient implemented in this manner predicts a partially asymmetric V-shaped LDF like the PM + RS and PM + OI models.

sumptions in order to account for these domain-specific features of spatial STM for serial order.

# References

Aaronson, D. (1968). Temporal course of perception in an immediate recall task. *Journal of Experimental Psychology, 76,* 129–140. http://dx.doi .org/10.1037/h0025290

Agam, Y., Bullock, D., & Sekuler, R. (2005). Imitating unfamiliar sequences of connected linear motions. *Journal of Neurophysiology, 94,* 2832–2843. http://dx.doi.org/10.1152/jn.00366.2005

Agam, Y., Galperin, H., Gold, B. J., & Sekuler, R. (2007). Learning to imitate novel motion sequences. *Journal of Vision, 7,* 1–17. http://dx.doi .org/10.1167/7.5.1

Agam, Y., Huang, J., & Sekuler, R. (2010). Neural correlates of sequence encoding in visuomotor learning. *Journal of Neurophysiology, 103,* 1418–1424. http://dx.doi.org/10.1152/jn.00662.2009

Agam, Y., & Sekuler, R. (2007). Interactions between working memory and visual perception: An ERP/EEG study. *NeuroImage, 36,* 933–942. http://dx.doi.org/10.1016/j.neuroimage.2007.04.014

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19,* 716–723.

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

Anderson, J. R., & Matessa, M. P. (1997). A production system theory of serial memory. *Psychological Review, 104,* 728–748. http://dx.doi.org/ 10.1037/0033-295X.104.4.728

Averbeck, B. B., Chafee, M. V., Crowe, D. A., & Georgopoulos, A. P. (2002). Parallel processing of serial movements in prefrontal cortex. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 99,* 13172–13177. http://dx.doi.org/10.1073/pnas .162485599

Averbeck, B. B., Chafee, M. V., Crowe, D. A., & Georgopoulos, A. P. (2003). Neural activity in prefrontal cortex during copying geometrical shapes I: Single cells encode shape, sequence, and metric parameters. *Experimental Brain Research, 150,* 127–141.

Averbeck, B. B., Crowe, D. A., Chafee, M. V., & Georgopoulos, A. P. (2003). Neural activity in prefrontal cortex during copying geometrical shapes II: Decoding shape segments from neural ensembles. *Experimental Brain Research, 150,* 142–153.

Avons, S. E. (1998). Serial report and item recognition of novel visual patterns. *British Journal of Psychology, 89,* 285–308. http://dx.doi.org/ 10.1111/j.2044-8295.1998.tb02685.x

Avons, S. E. (2007). Spatial span under translation: A study of reference frames. *Memory & Cognition, 35,* 402–417. http://dx.doi.org/10.3758/ BF03193281

Avons, S. E., & Mason, A. (1999). Effects of visual similarity on serial report and item recognition. *Quarterly Journal of Experimental Psychology, 52A,* 217–240. http://dx.doi.org/10.1080/027249899391296

Avons, S. E., & Oswald, K. (2008). Using temporal order to identify spatial reference frames. *Perception & Psychophysics, 70,* 1068–1080. http:// dx.doi.org/10.3758/PP.70.6.1068

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

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

Baddeley, A. D. (2007). *Working memory, thought and action.* Oxford, UK: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/ 9780198528012.001.0001

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105,* 158–173. http://dx.doi.org/10.1037//0033-295X.105.1.158

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47–89).

New York, NY: Academic Press. http://dx.doi.org/10.1016/S0079-7421(08)60452-1

Bernardis, P., & Shallice, T. (2011). Frames of reference in spatial span. *Quarterly Journal of Experimental Psychology, 64,* 2438–2454. http:// dx.doi.org/10.1080/17470218.2011.604202

Bohland, J. W., Bullock, D., & Guenther, F. H. (2009). Neural representations and mechanisms for the performance of simple speech sequences. *Journal of Cognitive Neuroscience, 7,* 1504–1529. http://dx.doi.org/ 10.1162/jocn.2009.21306

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

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

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

Bullock, D. (2004). Adaptive neural models of queuing and timing in fluent action. *Trends in Cognitive Sciences, 8,* 426–433. http://dx.doi.org/ 10.1016/j.tics.2004.07.003

Bullock, D., & Rhodes, B. (2003). Competitive queuing for serial planning and performance. In M. Arbib (Ed.), *Handbook of brain theory and neural networks* (2nd ed., pp. 241–244). Cambridge, MA: MIT Press.

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

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. http://dx.doi.org/10.1037//0033-295X.106.3.551

Burgess, N., & Hitch, G. J. (2005). Computational models of working memory: Putting long-term memory into context. *Trends in Cognitive Sciences, 9,* 535–541. http://dx.doi.org/10.1016/j.tics.2005.09.011

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodal inference: A practical information-theoretic approach.* New York, NY: Springer-Verlag.

Cooper, R. P., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology, 17,* 297–338. http://dx .doi.org/10.1080/026432900380427

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International, 34,* 891B.

Couture, M., & Tremblay, S. (2006). Exploring the characteristics of the visuospatial Hebb repetition effect. *Memory & Cognition, 34,* 1720–1729. http://dx.doi.org/10.3758/BF03195933

Davelaar, E. J. (2007). Sequential retrieval and inhibition of parallel (re)activated representations: A neurocomputational comparison of competitive queuing and resampling models. *Adaptive Behavior, 15,* 51–71. http://dx.doi.org/10.1177/1059712306076250

De Lillo, C. (2004). Imposing structure on a Corsi-type task: Evidence for hierarchical organization based on spatial proximity in serial-spatial memory. *Brain & Cognition, 55,* 415–426. http://dx.doi.org/10.1016/j .bandc.2004.02.071

De Lillo, C., & Lesk, V. (2010). Spatial clustering and hierarchical coding in immediate serial recall. *European Journal of Cognitive Psychology, 22,* 216–246. http://dx.doi.org/10.1080/09541440902757918

Della Sala, S., Gray, C., Baddeley, A. D., Allamano, N., & Wilson, L. (1999). Pattern span: A tool for unwielding visuo-spatial memory. *Neuropsychologia, 37,* 1189–1199. http://dx.doi.org/10.1016/S0028-3932(98)00159-6

Duncan, M., & Lewandowsky, S. (2005). The time course of response suppression: No evidence for a gradual release from inhibition. *Memory, 13,* 236–246. http://dx.doi.org/10.1080/09658210344000233

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

Farrand, P., Parmentier, F. B. R., & Jones, D. M. (2001). Temporal-spatial memory: Retrieval of spatial information does not reduce recency. *Acta Psychologica, 106,* 285–301. http://dx.doi.org/10.1016/S0001-6918(00)00054-8

Farrell, S. (2008). Multiple roles for time in short-term memory: Evidence from serial recall of order and timing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 128–145. http://dx.doi.org/10.1037/0278-7393.34.1.128

Farrell, S., Hurlstone, M. J., & Lewandowsky, S. (2012). Sequential dependencies in recall of sequences: Filling in the blanks. *Memory & Cognition, 41,* 938–952. http://dx.doi.org/10.3758/s13421-013-0310-0

Farrell, S., & Lelièvre, A. (2009). End anchoring in short-term order memory. *Journal of Memory and Language, 60,* 209–227. http://dx.doi.org/10.1016/j.jml.2008.09.004

Farrell, S., & Lelièvre, A. (2012). The dynamics of access to groups in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38,* 1659–1674.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic Bulletin & Review, 9,* 59–79. http://dx.doi.org/10.3758/BF03196257

Farrell, S., & Lewandowsky, S. (2004). Modeling transposition latencies: Constraints for theories of serial order memory. *Journal of Memory and Language, 51,* 115–135. http://dx.doi.org/10.1016/j.jml.2004.03.007

Farrell, S., & Lewandowsky, S. (2012). Response suppression contributes to recency in serial recall. *Memory & Cognition, 40,* 1070–1080. http://dx.doi.org/10.3758/s13421-012-0212-6

Farrell, S., & Oberauer, K. (2014). Working memory for cross-domain sequences. *Quarterly Journal of Experimental Psychology, 67,* 33–44. http://dx.doi.org/10.1080/17470218.2013.795177

Fischer, M. H. (2001). Probing spatial working memory with the Corsi Blocks Task. *Brain & Cognition, 45,* 143–154. http://dx.doi.org/10.1006/brcg.2000.1221

Frankish, C. (1985). Modality-specific grouping effects in short-term memory. *Journal of Memory and Language, 24,* 200–209. http://dx.doi.org/10.1016/0749-596X(85)90024-5

Frankish, C. (1989). Perceptual organization and precategorical acoustic storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 469–479. http://dx.doi.org/10.1037//0278-7393.15.3.469

Glasspool, D. W. (2005). Modelling serial order in behavior: Evidence from performance slips. In G. Houghton (Ed.), *Connectionist models in cognitive psychology* (pp. 241–270). Hove, UK: Psychology Press.

Glasspool, D. W., & Houghton, G. (2005). Serial order and consonant-vowel structure in a graphemic output buffer model. *Brain & Language, 94,* 304–330. http://dx.doi.org/10.1016/j.bandl.2005.01.006

Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps, and plans. In R. Rosen & F. M. Snell (Eds.), *Progress in theoretical biology* (Vol. 5, pp. 233–374). New York, NY: Academic Press.

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

Guérard, K., & Tremblay, S. (2008). Revisiting evidence for functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 556–569. http://dx.doi.org/10.1037/0278-7393.34.3.556

Guérard, K., & Tremblay, S., & Saint-Aubin, J. (2009). The processing of spatial information in short-term memory: Insights from eye tracking the path length effect. *Acta Psychologica, 13,* 452–457. http://dx.doi.org/10.1016/j.actpsy.2009.01.003

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

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

Henson, R. N. A. (1998b). Item repetition in short-term memory: Ranschburg repeated. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1162–1181. http://dx.doi.org/10.1037//0278-7393.24.5.1162

Henson, R. N. A. (1999). Positional information in short-term memory: Relative or absolute? *Memory & Cognition, 27,* 915–927. http://dx.doi.org/10.3758/BF03198544

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, 49*A, 80–115. http://dx.doi.org/10.1080/027249896392810

Hitch, G. J., Burgess, N., Towse, J. N., & Culpin, V. (1996). Temporal grouping effects in immediate recall: A working memory analysis. *Quarterly Journal of Experimental Psychology, 49*A, 116–139. http://dx.doi.org/10.1080/713755609

Hitch, G. J., Fastame, M. C., & Flude, B. (2005). How is the serial order of a verbal sequence coded? Some comparisons between models. *Memory, 13,* 247–258. http://dx.doi.org/10.1080/09658210344000314

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

Houghton, G. (2005). Introduction to connectionist models in cognitive psychology: Basic structures, processes and algorithms. In G. Houghton (Ed.), *Connectionist modelling in cognitive psychology* (pp. 1–42). Hove, UK: Psychology Press.

Hurlstone, M. J. (2010). *The problem of serial order in visuospatial short-term memory* (Unpublished doctoral thesis). University of York, York, UK.

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.

Jalbert, A., Saint-Aubin, J., & Tremblay, S. (2008). Visual similarity in short-term recall for where and when. *Quarterly Journal of Experimental Psychology, 61,* 353–360. http://dx.doi.org/10.1080/17470210701634537

Jones, D. M., Beaman, P., & Macken, W. J. (1996). The object-oriented episodic record model. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 209–238). London, UK: Erlbaum.

Jones, D. M., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1008–1018. http://dx.doi.org/10.1037//0278-7393.21.4.1008

Kemps, E. (1999). Effects of complexity on visuo-spatial working memory. *European Journal of Cognitive Psychology, 11,* 335–356. http://dx.doi.org/10.1080/713752320

Kemps, E. (2001). Complexity effects in visuo-spatial working memory: Implications for the role of long-term memory. *Memory, 9,* 13–27. http://dx.doi.org/10.1080/713752320

Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York, NY: Wiley.

Lawrence, B. M., Myerson, J., & Abrams, R. A. (2004). Interference with spatial working memory: An eye movement is more than a shift of attention. *Psychonomic Bulletin and Review, 11,* 488–494. http://dx.doi.org/10.3758/BF03196600

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

Lewandowsky, S., & Farrell, S. (2011). *Computational modeling in cognition: Principles and practice*. Thousand Oaks, CA: Sage.

Lewandowsky, S., & Murdock, B. B. (1989). Memory for serial order. *Psychological Review, 96,* 25–57. http://dx.doi.org/10.1037//0033-295X.101.3.539

Li, S.-C., Lewandowsky, S., & DeBrunner, V. E. (1996). Using parameter sensitivity and interdependence to predict model scope and falsifiability. *Journal of Experimental Psychology: General, 125,* 360–369. http://dx.doi.org/10.1037//0096-3445.125.4.360

Logie, R. (1995). *Visuo-spatial working memory*. Hove, UK: Erlbaum.

Maybery, M., Parmentier, F. B. R., & Jones, D. M. (2002). Grouping of list items reflected in the timing of recall: Implications for models of serial verbal memory. *Journal of Memory and Language, 47,* 360–385. http://dx.doi.org/10.1016/S0749-596X(02)00014-1

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7,* 308–313. http://dx.doi.org/10.1093/comjnl/7.4.308

Ng, L. H., & Maybery, M. T. (2002). Grouping in verbal short-term memory: Is position coded temporally? *Quarterly Journal of Experimental Psychology, 55A,* 391–424. http://dx.doi.org/10.1080/02724980143000343

Ng, L. H., & Maybery, M. T. (2005). Grouping in short-term memory: Do oscillators code the positions of items? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 175–181. http://dx.doi.org/10.1037/0278-7393.31.1.175

Oberauer, K. (2003). Understanding serial position curves in short-term recognition and recall. *Journal of Memory and Language, 49,* 469–483. http://dx.doi.org/10.1016/S0749-596X(03)00080-9

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

Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review, 110,* 683–712. http://dx.doi.org/10.1037/0033-295X.110.4.683

Paramentier, F. B. R. (2011). Exploring the determinants of memory for spatial sequences. In A. Vandierendonck & A. Smaleck (Eds.), *Spatial working memory*. Hove, UK: Psychology Press.

Parmentier, F. B. R., & Andrés, P. (2006). The impact of path crossing on visuo-spatial serial memory: Encoding or rehearsal effect? *Quarterly Journal of Experimental Psychology, 59,* 1867–1874. http://dx.doi.org/10.1080/17470210600872154

Parmentier, F. B. R., Andrés, P., Elford, G., & Jones, D. M. (2006). Organization of visuo-spatial serial memory: Interaction of temporal order with spatial and temporal grouping. *Psychological Research, 70,* 200–217. http://dx.doi.org/10.1007/s00426-004-0212-7

Parmentier, F. B. R., Elford, G., & Maybery, M. M. (2005). Transitional information in spatial serial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 412–427. http://dx.doi.org/10.1037/0278-7393.31.3.412

Parmentier, F. B. R., & Maybery, M. T. (2008). Equivalent effects of grouping by time, voice and location on response timing in verbal serial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1349–1355. http://dx.doi.org/10.1037/a0013258

Pearson, D. G., & Sahraie, A. (2003). Oculomotor control and maintenance of spatially and temporally distributed events in visuo-spatial working memory. *Quarterly Journal of Experimental Psychology, 56A,* 1089–1111.

Pitt, M. A., & Myung, J. M. (2002). When a good fit can be bad. *Trends in Cognitive Sciences, 6,* 421–425. http://dx.doi.org/10.1016/S1364-6613(02)01964-2

Rossi-Arnaud, C., Pieroni, L., & Baddeley, A. D. (2006). Symmetry and binding in visuo-spatial working memory. *Neuroscience, 139,* 393–400. http://dx.doi.org/10.1016/j.neuroscience.2005.10.048

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive motor performance. *Cognitive Science, 6,* 1–36. http://dx.doi.org/10.1016/S0364-0213(82)80004-9

Ryan, J. (1969a). Grouping and short-term memory: Different means and patterns of grouping. *Quarterly Journal of Experimental Psychology, 21*A, 137–147. http://dx.doi.org/10.1080/14640746908400206

Ryan, J. (1969b). Temporal grouping, rehearsal and short-term memory. *Quarterly Journal of Experimental Psychology, 21*A, 148–155. http://dx.doi.org/10.1080/14640746908400207

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6,* 461–464.

Smyth, M. M. (1996). Interference with rehearsal in spatial working memory in the absence of eye movements. *Quarterly Journal of Experimental Psychology, 40*A, 940–949. http://dx.doi.org/10.1080/713755669

Smyth, M. M., Hay, D. C., Hitch, G. J., & Horton, N. J. (2005). Serial position memory in the visual-spatial domain: Reconstructing sequences of unfamiliar faces. *Quarterly Journal of Experimental Psychology, 58*A, 909–930. http://dx.doi.org/10.1080/02724980443000412

Smyth, M. M., & Scholey, K. A. (1994). Interference in spatial immediate memory. *Memory & Cognition, 22,* 1–13. http://dx.doi.org/10.3758/BF03202756

Smyth, M. M., & Scholey, K. A. (1996). Serial order in spatial immediate memory. *Quarterly Journal of Experimental Psychology, 49*A, 159–177. http://dx.doi.org/10.1080/027249896392847

Solway, A., Murdock, B. B., & Kahana, M. J. (2012). Positional and temporal clustering in serial order memory. *Memory & Cognition, 20,* 177–190. http://dx.doi.org/10.3758/s13421-011-0142-8

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York, NY: Academic Press.

Thomas, J. G., Milner, H. R., & Haberlandt, K. F. (2003). Forward and backward recall: Different response time patterns, same retrieval order. *Psychological Science, 14,* 169–174. http://dx.doi.org/10.1111/1467-9280.01437

Tremblay, S., Guérard, K., Parmentier, F. B. R., Nicholls, A. P., & Jones, D. M. (2006). A spatial modality effect in serial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 1208–1215. http://dx.doi.org/10.1037/0278-7393.32.5.1208

Tremblay, S., Saint-Aubin, J., & Jalbert, A. (2006). Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. *Psychonomic Bulletin and Review, 13,* 452–457.

Vousden, J. I., & Brown, G. D. A. (1998). To repeat or not to repeat: The time course of response suppression in sequential behavior. In J. A. Bullinaria, D. W. Glasspool, & G. Houghton (Eds.), *Proceedings of the fourth neural computation and psychology workshop: Connectionist representations* (pp. 301–315). London, UK: Springer-Verlag.

Wagenmakers, E.-J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin & Review, 11,* 192–196. http://dx.doi.org/10.3758/BF03206482

Ward, G., Avons, S. E., & Melling, L. (2005). Serial position curves in short-term memory: Functional equivalence across modalities. *Memory, 13,* 308–317. http://dx.doi.org/10.1080/09658210344000279

*(Appendix follows)*

# Appendix

## Ancillary Simulations

Additional simulations were performed to determine whether the five models predict qualitatively similar LDFs to those shown in Figure 10D when they are able to reproduce the shape of the latency serial curve for ungrouped sequences depicted in Figure 8C. We generated model predictions using the same parameter values and modeling procedure used to generate the initial predictions shown in
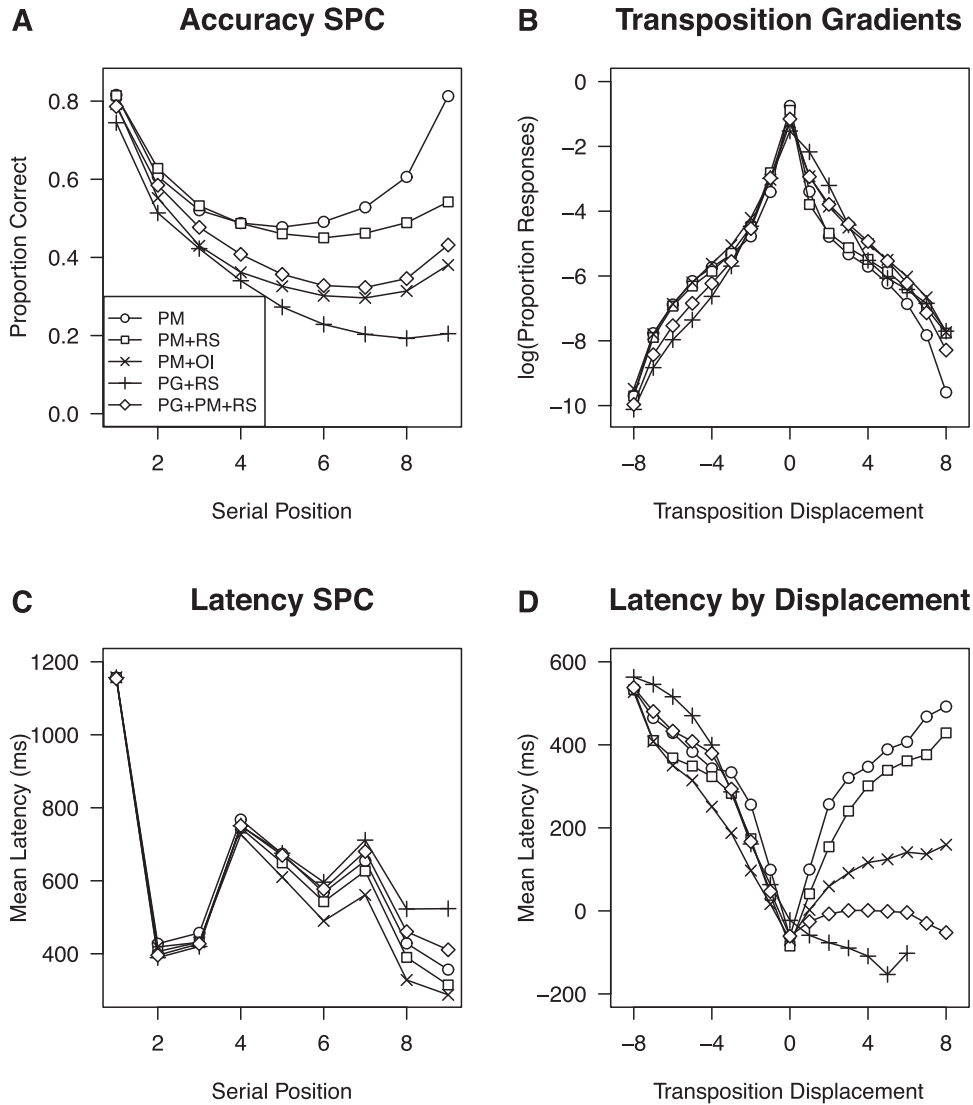


*Figure A1.* Response probability and recall latency predictions for five models of serial order. Panels show accuracy serial position curves (A), transposition gradients (B), latency serial position curves (C), and latency-displacement functions (D). PM = position marking; RS = response suppression; OI = output interference; PG = primacy gradient.

(*Appendix continues*)

Figure 4, but with the following noteworthy changes. First, the sequence length was increased from seven to nine items. Second, the standard deviation of noise σ parameter was increased from a value of .04 to .05 in order to bring overall levels of recall accuracy closer to those for ungrouped sequences in Experiment 3 (Figure 8A). Third, to reproduce the shape of the latency serial curve, on each trial, time constants were added to the predicted recall latencies for Output Positions 1, 4, 5, 6, and 7 for each model. The time constants were 40, 20, 15, 10, and 15 iterations, respectively. To introduce some between trial variability, on each simulation trial, the time constants were perturbed by random Gaussian noise ($SD = .05$) before they were added to the predicted latencies of the models. Finally, the iteration-to-ms scaling parameter $S$ was reduced from a value of 50 to 20 to bring the predicted latency serial position curves within the empirical range. All other aspects of the simulations were exactly the same as for those reported at the outset.

The predictions of the models can be inspected in Figure A1. It can be seen from inspection of Panel C that all models accurately reproduced the shape of the latency serial curve, but, critically, their predicted LDFs shown in Panel D do not differ qualitatively from those illustrated in Figure 9D.